



# Predicting the Carcinogenicity of Chemicals in Humans from Rodent Bioassay Data

## Citation

Goodman, Gay, and Richard Wilson. 1991. Predicting the carcinogenicity of chemicals in humans from rodent bioassay data. *Environmental Health Perspectives* 94: 195-218.

## Published Version

doi:10.2307/3431313

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:4459222>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Predicting the Carcinogenicity of Chemicals in Humans from Rodent Bioassay Data

by Gay Goodman\*<sup>†</sup> and Richard Wilson\*

Regulatory agencies currently rely on rodent carcinogenicity bioassay data to predict whether or not a given chemical poses a carcinogenic threat to humans. We argue that it is always more useful to know a chemical's carcinogenic potency (with confidence limits) than to be able to say only qualitatively that it has been found to be a carcinogen. In a typical bioassay, a chemical is administered to groups of 50 to 100 rodents at the highest feasible level (the maximum tolerated dose) and rarely at less than 1/10 this dose in order to maximize the statistical significance of any increase in tumors that might result. Recently, much experimental work has focused on the mechanisms by which site-specific toxicity arising from chronic administration at the maximum tolerated dose may lead to carcinogenicity. Extrapolation of high-dose results to low doses does not take into consideration the possibility of a threshold dose, below which the carcinogenic potency is much lower or even zero. Threshold dose-response phenomena may be much more relevant to the etiology of cancer in the rodent bioassays than was earlier realized; if so, there is an even greater need for establishing dose-dependent potency estimates. The emphasis of this review is on the interspecies comparison of high-dose potencies. The qualitative and quantitative comparison of carcinogenicities between mice and rats and between rodents and humans is reviewed and discussed. We conclude that there is a good qualitative (yes/no) correlation for both the rat/mouse and the rodent/human comparison. There is also a good correlation of the carcinogenic potencies between rats and mice, and the upper limits on potencies in humans are consistent with rodent potencies for those chemicals for which human exposure data are available. For the rodent/human comparison, the best estimate of the interspecies potency factor is lognormally distributed around 1 when the potencies in both species are measured in units of (mg/kg-day)<sup>-1</sup>.

## Introduction

Whenever researchers have thought that there is a simplifying feature about cancer, further studies have arrived to demonstrate complications. This particularly applies to any attempts to develop rules for discovering potential human carcinogens and for establishing low-risk exposure limits. Of course, we do not deliberately expose people to a chemical merely to find out whether it causes cancer; human epidemiological information comes from accidental, incidental, or therapeutic exposures. This forces us to search for other methods of predicting whether a chemical is carcinogenic to humans. Although the development of short-term *in vitro* and *in vivo* tests holds promise for the future in predicting the carcinogenicity of substances, at the very least we are left with the task of interpreting the existing rodent bioassay data. The problem then becomes: how should we utilize the results of carcinogenicity tests in rodents for predicting cancer risk in humans?

In long-term bioassays, chemicals are tested at the highest possible dose in order to maximize the sensitivity for detecting carcinogenic effects. To take a hypothetical example, if a carcinogen induces a particular type of neoplasm with frequency

0.2 at dose  $d$  and with frequency 0.02 when the dose is  $(1/10)d$ , then in a test of 100 animals the carcinogenicity of the chemical would be detectable above background at the higher but not the lower dose. The maximum dose tested (MaxD) is usually set at the maximum tolerated dose (MTD), the highest dose that does not cause death, organ toxicity, or severe weight loss during a chronic dosing regime of several months' duration. The other doses in the bioassay are usually chosen as some fraction of the MTD, and are only rarely as small or smaller than MTD/10. One issue surrounding the fact that the bioassays are conducted at high doses is that the shape of the dose-response curve at much lower doses is not necessarily predicted by the dose-response near the MTD. Although we return to this point when discussing possible mechanisms of chemical carcinogenesis, the high-dose to low-dose extrapolation is not the subject of this review. It is important to acknowledge that there are two separate aspects of determining human risk from data taken in rodent bioassays, namely *a*) the high-dose/low-dose extrapolation and *b*) the interspecies extrapolation of high-dose potencies. We emphasize that here we are addressing mainly the second aspect, in which carcinogenic potencies obtained in rodents are used to predict potencies in humans exposed at similarly high doses.

## Which Chemicals Are Carcinogens?

Twenty years ago, there were not many agents that were

\*Department of Physics, Harvard University, Cambridge, MA 02138.

<sup>†</sup>School of Public Health, Boston, MA 02115.

Address reprint requests to G. Goodman, Gradient Corporation, 44 Brattle Street, Cambridge, MA 02138.

known to cause cancer in either laboratory animals or man. It was considered prudent to impose strict regulatory restrictions on the use of all such chemicals. It was generally anticipated that few new carcinogens would be discovered; certainly the percentage of chemicals tested for carcinogenicity that turned up "positive" was low at that time. In one study of 120 pesticides and industrial chemicals given orally for 18 months at the MTD, only 11 of the chemicals tested (9%) caused cancer in mice (1). This oft-quoted study has been cited as evidence that chemicals can be given at near-toxic doses without automatically triggering carcinogenicity secondary to toxic effects (2). Although we agree for other reasons (see above) that it is reasonable to test at the MTD, the results of Innes et al. (1) bear further analysis. In fact, among the 109 chemicals not found to be carcinogenic, 20 (17%) were considered by the authors to require further investigation. Ames and Gold (3) stated that this study was less thorough than modern protocols and pointed out that only one species was tested. We also note that even with the identical protocol, some additional chemicals would most likely have produced tumors if the duration of exposure or even observation had been extended to 24 months.

In the intervening years, the available information has increased. We now know of hundreds of chemicals that can cause cancer in animals; many are impossible to ban, and some are of natural origin (4). New cancer bioassays involve dosing animals daily for a period on the order of a lifetime: 2 years for rats and mice. More animals are used, both sexes of two species are exposed, generally higher doses are administered, and the effects at several dose levels (usually at least three) are included (5). Of 266 chemicals considered to have been adequately tested in rats and mice of both sexes at or near the MTD in 2-year protocols by the National Toxicology Program (NTP) and the National Cancer Institute (NCI), 51 percent were found to be positive for carcinogenicity in at least one sex-species group; the evaluation of positivity was according to the conclusions stated in the NCI/NTP Technical Reports (6,7).

Almost the same data has been looked at in a slightly different way by Byrd et al. (8). Given that all experiments consisted of at least two dose groups and one control group, their criterion for positivity was an increased site-specific tumor incidence with  $p < 0.025$  for an individual dose group, or  $p < 0.05$  for any two dose groups (if one or both failed to meet the more stringent criterion), by Fisher exact test. Out of a total of 290 chemicals tested by the NCI/NTP and included on their Carcinogenic Bioassay Database System (CBDS) magnetic tapes as of December, 1987, 50% were positive in mice, 52% were positive in rats, and 71% were positive in at least one sex-species group. If the tumor-bearing male and female animals were combined before testing for statistical significance, then the fraction labeled positive would be greater still. We note also that because the number of animals per dose group is usually 50 or 100, tumor incidences below the level of 5% may go unobserved or be discounted due to lack of statistical significance. If the sensitivity were improved, for example, by using 1000 animals per dose group, many more chemicals might be shown to be carcinogenic. However, since the statistical criteria for positivity used by Byrd et al. were rather liberal (8), it is likely that many of the statistically significant differences in tumor incidence were false positives, reflecting merely random variability (9-11).

Along with an increase in the power of animal tests to identify carcinogenic substances, there has been a parallel improvement in the methodology for detecting trace amounts of chemicals. Thirty years ago, an upper limit on sensitivity of a few parts per million was common. It seemed likely that any carcinogen found in foodstuffs at that level would pose a significant risk and remedial action should be demanded. So few chemicals in commercial use were known to be carcinogens that it seemed that they could be avoided and replaced by alternatives. This was the rationale that led to the inclusion of the Delaney Clause in the Food Additive Amendments of 1958 [reviewed by Hutt (12)]. If a few turned out to be essential or in unusual demand, then special legislation could be enacted, as later happened for saccharin. But we can now detect trace contaminant levels to one part in  $10^{12}$  in some cases. As a result, in typical drinking water supplies, at least 20 chemicals known to be rodent carcinogens are well above the detection level, and many other chemicals are present for which information as to carcinogenicity is either unavailable or inadequate. Because we find it technically or economically unfeasible to remove all such chemicals from our drinking water, our diet, and the environment in general, it becomes necessary to rank chemicals according to the risk they are expected to pose to human health. One means of achieving such a ranking is by measuring carcinogenic potencies in animals and finding some means of extrapolating these results quantitatively to humans.

It seems therefore unhelpful to classify chemicals as carcinogens or noncarcinogens, but to assume that all chemicals are carcinogenic and that some have too low a potency to produce a statistically significant increase in tumors with a given experimental protocol. This is a variation on a statement by Paracelsus: "All things are poisons, for there is nothing without poisonous qualities. It is only the dose which makes a thing a poison." [quoted by the Scientific Committee of the Food Safety Council (13)]. Under this scheme, a chemical that is truly noncarcinogenic at any dose is a special case and would be called a carcinogen of zero potency.

The assumption that all chemicals may be carcinogenic only leads to a useful procedure if one can determine the quantitative carcinogenic potency for each chemical and an upper limit to the potency in cases where the potency is too low to be detected. In the present paper we review the work of our group and of others on the interspecies correlation of carcinogenic potencies, particularly between mice and rats. Further analysis is provided of the methodology, efficacy, and limitations for the use of rodent bioassay data in predicting human carcinogenic potencies. Gregory (14) and Dybing (15) have recently reviewed interspecies comparisons with respect to predicting carcinogenic risk in humans; we have attempted not to duplicate their efforts, and the reader is referred to these papers for alternative points of view.

## Quantitation: Relating Carcinogenic Potency to Other Parameters

Currently, there exists no good alternative to animal bioassays for obtaining quantitative information on the carcinogenic potential of chemicals in any species. The animal bioassay data base contains information on the tumorigenic effects of almost 1000

chemicals. Although many chemicals have been tested for their ability to induce point mutations in *Salmonella* strains, and some have been assayed for other genotoxic, cytotoxic, or preneoplastic effects in a variety of short-term *in vitro* and *in vivo* tests, a quantitative method that begins with data derived from such tests and ends with a prediction of carcinogenic potency in humans has not yet been put forward. As knowledge of the molecular mechanisms of action of genotoxic carcinogens increases, short-term tests for genotoxicity should be designed to reflect the importance of different genetic end points in primary carcinogenesis (16). Research into the mechanisms of action of nongenotoxic carcinogens, along with novel experiments to determine their likely importance in animals or humans when moderate levels of endogenous or environmental initiators are present simultaneously, should point the way toward development of yet another generation of short-term tests. For example, knowledge that the tumor promoting action of 12-*O*-tetradecanoylphorbol-13-acetate (TPA) is mediated by activation of protein kinase C, which is a key component of transmembrane signal transduction relevant to the regulation of cell division, suggests that other tumor promoters might function by disrupting the same or related pathways of signal transduction (17). It is conceivable that a simple *in vitro* test for promoters having such activity could be designed (18).

Although there is good reason to be optimistic that short-term tests will someday provide the means for predicting the potency of human carcinogens, that day has not yet arrived (19). It is clear that positivity on short-term tests for genotoxicity is highly correlated with carcinogenicity for the known human carcinogens: of the 23 chemicals and chemical combinations designated by the International Agency for Research on Cancer (IARC) to be causally associated with human cancer, all but two (asbestos and conjugated estrogens) are genotoxic or are expected to be so on the basis of chemical structure (20). But because of the paucity of data on human carcinogenicity and on definitive human non-carcinogens, there is only a very limited opportunity for attempting to calibrate short-term test results to these data. The emphasis has thus been on demonstrating agreement between short-term tests and rodent carcinogenicity bioassays, both qualitative (21-25) and quantitative (26-30). When the specificity of existing short-term tests at distinguishing between rodent noncarcinogens and carcinogens is analyzed, their power as predictive tools is less impressive than when sensitivity to rodent carcinogens is the only criterion (31,32). The finding that a large proportion of known rodent carcinogens are not genotoxic in short-term tests (23) underscores the necessity of developing short-term tests for nongenotoxic carcinogens. Since the analysis of the efficacy of all short-term tests is focused primarily on existing carcinogenicity data obtained in long-term animal studies, it is important to try and ascertain the best methods for retrieving information relevant to human carcinogenic potency from the bioassay data and to quantify the expected uncertainty in both the upper and lower bounds of such potency estimates.

## Dose Conventions

Before testing for carcinogenicity was even contemplated, the testing of chemicals for toxic effects in animals was used for predicting toxic effects in humans. There are several ways to formulate the dosage administered for the purposes of such a

comparison; see Calabrese (33) for a comprehensive review. The usual convention for normalizing dose is on a weight/weight basis; weight of the chemical divided by the weight of the animal. The reporting of dose as a fraction of body weight does not imply that equal effects will be produced in two species if the dosage is scaled according to their weights. It is merely a convention and no more. However, this convention arose out of an expectation that the amount of chemical (expressed as a weight)  $w_h$  required to achieve the same toxic effect in the heavier animal that is produced in the lighter animal by the amount  $w_l$  is  $w_h = w_l(W_h/W_l)$ , where  $W_h$  and  $W_l$  are the weights of the heavier and lighter animals, respectively. A crude surface area correction is sometimes applied to the usual dose convention, such that the prediction for equal toxicity becomes  $w_h = w_l(W_h/W_l)^{2/3}$ . If instead of the weight of chemical we now use the convention in which dose is normalized on a weight/weight basis, i.e., dose  $d = w/W$ , then the surface-area-corrected prediction for equal toxicities becomes  $d_h = d_l(W_h/W_l)^{1/3}$ . An excellent interspecies correlation of minimal toxic doses was found for a series of 18 chemotherapeutic agents when the scaling factor was based on body surface area (34). It was later realized that these chemicals are not readily metabolized by liver microsomal enzymes, and thus interspecies variation in such metabolism was not a major factor in the comparison (35). Calabrese (33) concludes that in the absence of an efficient drug-metabolizing system, a drug's toxic effects are similar among species when the dose is scaled on the basis of surface area. However, another study of the acute toxicity of 400 chemicals revealed a good correlation with body weight (and not body weight to the 2/3 power) for more than 80% of the chemicals (36). One might expect that for all chemicals within a given chemical class, the same dose convention would be appropriate to the interspecies comparison of toxic effects.

For carcinogenic potency, there exists a similar uncertainty as to which dose convention to use to simplify the interspecies comparison. In addition, since cancer typically develops as a result of long-term exposure to a carcinogenic substance, the factor of time enters the determination of potency. With daily administration, a steady-state concentration of the active agent in the animal's system is approached. There is no obvious reason that the same interspecies scaling factor should apply here as is relevant to acute effects such as the response to a single dose of a drug. In the present paper, for simplicity, we refer to the daily dose administered on a weight/weight basis (typically milligrams per kilogram) for the lifetime of the animal or human. Use of this convention, by analogy with that described above for acute toxicity, does not imply that weight scaling always produces a 1:1 correspondence between effective carcinogenic doses in two species. For example, B6C3F<sub>1</sub> mice and F344 rats demonstrated a better correlation of carcinogenic potencies with surface area scaling than with weight scaling, but for the same (B6C3F<sub>1</sub>) mice strain and a different (Osborne-Mendel) rat strain the opposite was true: weight scaling yielded a higher degree of potency correlation than did surface area scaling (37). Based upon recent work on physiological parameters controlling pharmacokinetics, Travis et al. have suggested that the optimum interspecies scaling factor (for this chemical and this organ) is in fact closer to weight to the  $\frac{3}{4}$  power (38). We consider that the best estimate for an interspecies scaling factor is an unknown, to be determined from existing data. In the risk-assessment procedure the most

accurate scaling factor ought, in principal, to be adopted on a case-by-case basis; unfortunately, relevant information is only rarely available.

## Indexes of Carcinogenic Potency

The data from long-term bioassays can be analyzed according to any one of a number of dose-response models resulting, for each experiment, in a measure of the chemical's carcinogenic potency. One standard index of carcinogenic potency, is  $\beta$ , which here we take to be the slope of the dose-response relationship at the dose of interest (37,39). In principle, the potency can be a function of dose  $d$ :  $\beta = \beta(d)$ . In the limit where  $\beta$  is constant, a formula can be derived for the fractional probability  $P$  of producing at least one tumor at a given site:

$$P(d) = 1 - (1 - \alpha) \exp\left(\frac{-\beta d}{1 - \alpha}\right) \quad (1)$$

where  $\alpha$  is the background rate of animals with tumors at that site. This formula was specifically constructed to reach but not exceed  $P = 1$  at high doses and to have the (linear) low-dose limit  $P = \alpha + \beta d$ , which means that  $\beta$  is the initial slope of  $P$  versus  $d$  (39). The model was originally derived in terms of tumor-bearing animals rather than animals with tumors at a specific site, but the form is the same in either case.

Useful bioassay data fall approximately in the range of 10 to 95% tumor incidence at the most sensitive site. Gold and co-workers have chosen an expression for potency that emphasizes the region where data are available. Their index of potency is  $1/\text{TD}_{50}$ , where  $\text{TD}_{50}$  is defined as the dose that, if administered daily for the "standard lifespan" of the species, "will halve the mortality-corrected estimate of the probability of remaining tumorless [at a given site] throughout that period" (40). To derive this index, they used a mathematical model which, like Eq. (1), assumes linearity at low doses. Indeed, in the absence of intercurrent mortality (i.e., deaths during the term of the study), the parameter  $\beta$  is equivalent to  $(\ln 2)/\text{TD}_{50}$  (40). Because  $1/\text{TD}_{50}$  is defined in terms of a dose, the  $\text{TD}_{50}$ , which falls in the middle of the range for which there are useful data, is fairly independent of the model used in its derivation. In most cases the  $\text{TD}_{50}$  turns out to be close in value to the maximum administered dose, which approximates the MTD (41-43).

## Extrapolation from High-Dose to Low-Dose Potencies

A major problem with testing at doses near the MTD is that some toxic effects may be inevitable. Chemicals that are carcinogenic only as a result of target-organ toxicity might exhibit a dose-response relationship that reflects the secondary nature of their carcinogenicity. Since toxic effects usually appear only above a threshold dose, one might anticipate that such secondary carcinogens would demonstrate a similar threshold behavior for tumorigenicity. Other chemicals might be primary carcinogens, and it is generally assumed that they would have some quantifiable carcinogenic potential at any dose. Given that such a distinction is meaningful, both classes of animal carcinogens are likely to be relevant in principle to human cancer, even if only the primary carcinogens produce cancer at low exposure levels in

humans. (We operationally define a "low" dose of a given substance rather loosely: a dose that is much less than the MTD for that substance in the species being addressed.)

When attempting to extrapolate from a high-dose measure of potency (such as  $\beta$  or  $1/\text{TD}_{50}$ ) to a low-dose potency, the choice of dose-response model becomes crucial and can shift the outcome by many orders of magnitude. In some sense then, use of an index of carcinogenic potency that only reflects experimentally measurable (high-dose) effects permits one to beg the question of the validity of any given method for extrapolating from high to low doses. We are aware that there is no single correct way to determine low-dose potencies. The problem becomes a serious one in cases where there exists a threshold dose for carcinogenicity. But for most if not all genotoxic carcinogens that have been tested adequately at low doses, a low-dose threshold has not been found. For such chemicals, the assumption of a linear dose response down to zero dose provides a reasonable estimate of the potency at low doses for the most sensitive site. Examples are 2-acetylaminofluorene (44,45), diethylnitrosamine and dimethylnitrosamine (46), and *N*-butyl-*N*-(4-hydroxybutyl)-nitrosamine (47).

There is now a substantial amount of evidence from animal experiments indicating that cell proliferation is sometimes associated with the process of chemical carcinogenesis when chemicals are administered at doses near the MTD. The degree of local cell proliferation is apparently a good index for cell killing; indeed, it has been suggested by Swenberg (48) that, whereas the MTD as currently determined often results in a 20-fold increase in cell proliferation in some organ, the MTD ought to be redefined as that dose which results in no more than a 3- or 4-fold increase in cell proliferation in any organ.

The possibility now exists that threshold phenomena for carcinogenicity may be explained by increased cell proliferation. In rats given 2-acetylaminofluorene (2-AAF), which has a hockey-stick-shaped dose response for mouse bladder tumors with the number of tumors increasing sharply above 60 ppm, bladder hyperplasia also increases nonlinearly above 60 ppm (49). Similarly, investigation of the tumorigenicity of saccharin revealed that the threshold dose for bladder tumors in the male rat is explained by the dose-response for precipitation of silicates in the urine; female rats do not form these silicates in response to saccharin and are likewise not induced to form bladder tumors (49). A possible implication of findings such as these is that for chemicals that produce tumors in a given organ by a mechanism unrelated to genotoxicity (as evidenced by failure to produce genotoxic effects in short-term tests and either absence of DNA adduct formation in the target organ or a threshold response for tumorigenesis that does not match the dose-response for DNA adduction), induction of cell proliferation may be a necessary condition for tumorigenesis. But what we do not yet know with certainty is whether the isolated cases in which the dose-response for cell proliferation has been measured is indicative of the pattern one would find for most or all nongenotoxic chemicals: i.e., that there is a threshold dose below which a chemical produces neither cell proliferation nor neoplastic growth. It is vital that such dose-response information be forthcoming if we are to gain any measure of confidence in our ability to predict carcinogenic potencies based on cell proliferation studies.

It is also far from certain that most genotoxic agents that cause cancer do so via genotoxic mechanisms, either in rodents or in humans. We have reported elsewhere that the carcinogenic potency (as obtained from the rodent bioassay) is more strongly related to the MTD for nonmutagens than for mutagens, although the differences are small (50). Our findings are consistent with the hypothesis that, even for most mutagens, at high doses carcinogenicity is induced via mechanisms associated with toxicity. Similarly, for genotoxic chemicals known to cause cancer in humans, most epidemiological evidence comes from high-dose exposures, approaching the MTD in many cases. Probably the best-studied human carcinogen is tobacco smoke, which produces acute toxic effects in the lungs and respiratory system at all levels of usage. It may be argued that the target tissue dose level is high for the duration of inhalation, regardless of how few or how many cigarettes are smoked per day. For this reason, toxic effects cannot be ruled out as a contributing cause or even as the main cause of smoking-related carcinogenesis, despite the fact that cigarette smoke contains potent genotoxins.

Although we have touched briefly upon the low-dose extrapolation problem, it is not the subject of the present paper. Readers are referred to the review of Zeise et al. (51) for a thorough discussion of dose response based on available low-dose data. In the following, interspecies comparisons refer primarily to potencies obtained when the exposure doses were at high or intermediate levels for each of the species under consideration.

## Interspecies Conversion Factor

In this review, we extrapolate the potency from species *a* to species *b* via the relation  $\beta_b = K_{ba}\beta_a$  (52). This equation defines the factor  $K_{ba}$ , which is termed an interspecies conversion factor. Because of a multitude of pharmacokinetic parameters determining a given chemical's effective dose and the species-variability of particular biochemical pathways which decide its actual carcinogenicity,  $K$  is not expected to have the same value for all chemicals. For the comparison of carcinogenic potencies between mice and rats, it has been found that  $K$  varies a factor of 20 over more than 200 rodent carcinogens (37,53). Our task is to find appropriate values for the factor  $K_{hr}$  for the rodent to human conversion. (Here we use subscript *R* to indicate rodents in general, reserving the lowercase *r* for use later when we refer specifically to rats.) We begin with the hypothesis that a given chemical or agent that is a carcinogen in one rodent species is carcinogenic in another. We further propose that the carcinogenic potency in humans is close to that in rodents ( $K_{hr} \approx 1$ ) when the potencies are measured in units of  $(\text{mg/kg-day})^{-1}$ , and how close  $K_{hr}$  is to 1 can be derived from experimental and epidemiological data. A scientific study may disprove this proposition for a given chemical or class of chemicals, in which case our hypothesis would have to be abandoned or modified. Possible modification might be that  $K_{hr}$  is, on average, closer to 10, or to 0.1. The whole process of risk assessment should therefore be iterative. As biological understanding improves upon the assumptions, the assessment may be modified appropriately. Some government agencies have made various assumptions about this factor in their risk assessments; these are summarized in Table 1. But we reemphasize that  $K_{hr}$  is not a constant; it varies from chemical to chemical, although generalizations made on the basis of chemical class could be warranted.

Table 1. Comparison of relative interspecies potency factors used by some policy makers.<sup>a</sup>

Method	Rat	Mouse	Human	Reference
mg/kg-day (chemical weight/ body weight)	1	1	1	Crouch and Wilson (39); Crouch (37) (Food and Drug Administration)
Percentage of diet (chemical weight/ diet weight)	1	0.38	1.8	Anderson et al. (III)
(mg/kg-day) $\times (W_r/W_x)^{1/4}$ (surface area correction)	1	0.44	5.9	Anderson et al. (III) (Environmental Protec- tion Agency)
(mg/kg-day) $\times (W_r/W_x)^{1/4}$	1	0.29	14	Travis (38)
mg/kg-day $\times$ lifetime	1	1	35	NAS (162)

<sup>a</sup>The numbers given are relative to potency in rats = 1.  $W_x$  is the weight of species *x* (mouse or human);  $W_r$  is the weight of the rat. Weights used in the calculations were rat: 0.35 kg, mouse: 0.03 kg, human: 70 kg.

## Deriving Potency in Humans from Potency in Rodents

There are several possible procedures for deriving carcinogenic potency in humans from the carcinogenic potency in rodents. *a*) Understanding the biological processes relevant to chemical carcinogenicity, their kinetics, and how these differ between rodents and humans. *b*) Comparing "natural," background tumorigenesis in rodents and humans. *c*) Comparing carcinogenicity of chemicals in rodents and humans in those cases where there is adequate epidemiology in humans. *d*) Comparing carcinogenicity of chemicals in rats and mice, thereby obtaining what is probably an upper limit on the accuracy of the comparison between rodents and humans.

## Understanding Biological Mechanisms of Carcinogenesis

The fact that interspecies variation in biological processes plays an enormous part in carcinogenesis can be seen by a *reductio-ad-absurdum* argument due to Peto (54). Suppose that all tissues of a human and a mouse were equally likely to develop neoplasms when, for example, bombarded uniformly by ionizing radiation. The probability of getting a tumor in any given time period is therefore greater for a human than a mouse (simply because there is more tissue) by the ratio  $M_h/M_m$ , where  $M_h$  and  $M_m$  are the masses of the human and the mouse respectively. Now the incidence of cancer is well known to increase with age by a power of the time (55-57), the exact power depending upon the type of malignancy. Let us assume the fourth power. Then the probability of developing a tumor during the mouse's lifetime  $T_m$  is less than that of developing one during the human's lifetime  $T_h$  by  $(T_m/T_h)^4$ . Therefore we expect that the lifetime incidence of radiation-induced cancer in mice will be less than that in humans by a factor of approximately

$$(M_m/M_h)(T_m/T_h)^4 \approx (30 \text{ g}/70 \text{ kg})(2 \text{ years}/70 \text{ years})^4 \approx 1/10^9.$$

Yet evidence suggests that the actual mouse/human ratio for radiation-induced cancer is closer to unity (58). We all know the answer to this paradox in general terms: it is because the sequence of biochemical and physiological events leading to carcinogenesis proceeds more rapidly in mice than in humans. The



rates of cell division are faster in smaller animals. The cardiac output of a mouse is 100% of its total blood volume per minute, whereas that of a man is only 5% of his blood volume per minute (59). These and other factors, both understood and unknown, result in differential probabilities and lag times for tumor induction, and different dose responses.

The problem facing us is to find the best estimate that we can of the actual potency ratio for chemical carcinogenesis between experimental animals and humans. Clearly, when estimating a given chemical's carcinogenic potency in humans from its potency in rodents, all available information on differences in the rates and mechanisms of its absorption, distribution to target organs, biochemical transformation to active or inactive metabolites, clearance, and mechanisms of carcinogenicity should be taken into consideration (60). Attention to differential pharmacokinetics may be of especial importance to the low-dose extrapolation (61). Unfortunately for the field of quantitative risk assessment, such information is lacking for most chemicals proven to be carcinogenic in animal bioassays. But even given detailed knowledge of pharmacokinetic and other relevant biological behavior in one species, there is not yet any formula for calculating quantitatively a chemical's tumorigenicity from its actions at the molecular, cellular, or organ level. On the other hand, where such data are available in two animal species for a chemical that has been tested for carcinogenicity in both, it should be possible to calibrate some biological parameters to carcinogenic potency. If these same parameters can be estimated in humans, for example, by measuring urine metabolites in exposed populations (62) or by analogy with related compounds (especially anticancer drugs) for which such data exists, then the animal-human extrapolation would thereby be facilitated.

The question of whether target organ toxicity is a precursor to tumorigenesis was recently addressed by Hoel and co-workers. Ninety-nine chemicals were tested for carcinogenicity and site-specific toxicity in long-term rodent bioassays; only 7 of the 53 which were positive for carcinogenicity exhibited target organ toxicity that could have been causal to the observed neoplasms (63). However, some local responses normally associated with toxicity, such as hyperplasia and inflammation, were considered by these authors to be corollaries of neoplasia rather than toxicity. By defining these responses as preneoplastic, it appears to us that they have ruled out *a priori* what might be an interesting effect. It would have been more informative to ask, instead: which toxic effects are associated with neoplasia? Therefore it is not clear to us whether they have adequately tested the relationship between toxicity and carcinogenicity.

Great strides have been made in the area of biologically based modeling of the rates of progress through the stages of carcinogenesis. Models in which carcinogenesis proceeds in accordance with two rate-limiting, heritable, cellular transitions (from normal to initiated and from initiated to transformed), and in which the clonal population size and the mitotic rate of the initiated cells are included among the variables, have been successful in explaining the incidence of various human cancers (64,65) and chemically induced bladder cancer in rats (66). Furthermore, experimentation based upon such models has permitted elucidation of the mechanism of action of at least one nongenotoxic, low-potency carcinogen: it was concluded that the tumor-promoting attributes of sodium saccharin are due to increases in the initiated

cell population and mitotic rate (but not the probability of initiation), and that these effects are secondary to cytotoxicity (67,68). Such advances in understanding the cellular phenomenology of carcinogenesis should result in the development of sensible, short-term, *in vivo* animal tests that might predict low-dose carcinogenic potency in humans with a higher degree of accuracy and at substantially less cost than a lifetime bioassay. There is also a need for studies designed to determine the low-dose response to tumor promoters in long-term animal bioassays in which the animals have been pretreated with an initiating agent (69). Such studies might then provide a calibration scale for past and future short-term *in vivo* initiation/promotion experiments, such as those proposed by Potter (70,71) or by Kunz et al. (72).

But for the time being, it must be kept in mind that the existing methods for modeling do not predict the carcinogenic potencies of chemicals; rather, they offer a means of predicting the number of tumors produced and their rate of appearance given a fixed set of variables such as the number of normal and initiated cells and the rates of mitosis, differentiation, and death. Many of these variables are expected to depend upon the particular chemical under test, the dose at which it is administered, and the background of initiated cells in the organ under study. We do not know of any published report in which carcinogenic potency was derived from an experimental system designed around a biologically based model, and which is directly applicable to interspecies extrapolation. We anticipate that future attempts will entail calculation of an interspecies potency factor for each class of chemical carcinogens, based upon the physiological parameters discussed above.

Our present inability to predict carcinogenic potency given detailed pharmacokinetic data and information as to genotoxicity on short-term tests reflects a general lack of understanding of the mechanisms of carcinogenesis. Amid much enthusiasm over the association of *ras* oncogenes with neoplastic cells, there has been a question as to whether mutation of a *ras* proto-oncogene is causative for the development of some human malignancies or is a consequence of them (73). This problem may have been solved by recent analysis of *ras* mutations in the leukemic cells of relapsed patients who had undergone chemotherapy. Four patients in whom mutations in *ras* genes were detected upon presentation no longer exhibited *ras* mutations upon relapse following chemotherapy-induced remission. It now appears that the bone marrow precursors of the leukemic white cells did not contain the *ras* mutation, which therefore must have arisen in a later stage of the disease (74). Other theories continue to be formulated and revised. The hypothesis that all carcinogens are mutagens (21) has been replaced by one which suggests that all initiators are genotoxic (75,76) and that tumor promoters are agents which directly or indirectly increase the mitotic rate of the initiated clonal cell population (66,64). It has been proposed that the carcinogenicity of nongenotoxins, including those that act only as tumor promoters or incomplete carcinogens, is secondary to *a*) site-specific cellular toxicity (4,77), perhaps mediated by oxygen radicals (78) or *b*) hormonal or immunological effects, including stimulation of cell proliferation.

The phenomenon of tumor progression is still not well understood, although it appears that this, at least for some cancers, may be a stage affected by environmental carcinogens. As discussed by Higginson, latent carcinoma of the prostate occurs with the same high (10%) incidence in 75-year-old black and white men in the U.S. and in Japanese men, whereas active prostatic

carcinoma which is then diagnosed is much rarer, occurring in the ratio 60:30:1 in these same populations (79). (For 1984–1986, the average yearly age-adjusted death rate for prostate cancer was 0.0232% in U.S. males and 0.0051% in Japanese males.) Differences in the incidence of the active form are thought to reflect environmental (perhaps dietary) influences on the rate of tumor progression. Japanese immigrants to the U.S. have a markedly higher death rate from cancer of the prostate than those remaining in Japan (80).

Heritable changes in DNA may only be among the first (and last) of a sequence of events necessary to the creation of a typical cancerous cell. There is even a question as to whether single-base mutation, which is the lesion detected in the typical bacterial mutagenesis assay, is on the pathway for creation of most neoplasms affecting internal organs. It has been pointed out that persons with the inherited syndrome xeroderma pigmentosum (XP), who suffer from one of several gene defects resulting in lack of an enzyme activity necessary for excision repair of mismatched basepairs, do not appear to have an increased incidence of internal cancers (81). This in spite of the fact that XP patients are more than 1000 times likelier to develop UV-induced skin cancer on exposed areas of the body than normal persons. Conversely, persons with Bloom's syndrome, an inherited disorder involving chromosomal fragility (resulting in frequent rearrangements and other aberrations), have an extraordinarily high age-adjusted death rate due to various types of cancer (81). Out of 103 Bloom's syndrome patients diagnosed in a 30-year period, 28 malignant neoplasms were detected at a mean age of 20.7 years (82). Cairns has concluded that simple basepair mutations are not likely to be the rate-limiting components in most human carcinogenesis, and that major genomic changes such as rearrangements and deletions are probably more important (81). It has also been noted that the presence of mutagenic substances in the urine of smokers (62) "results in only a moderate excess of urinary and pancreatic cancers and in no large excess of leukemias, lymphomas, or solid cancers at other sites distant from the respiratory tract" (83).

Accordingly, since we are as yet without a method for deriving carcinogenic potencies from biological first principles, it is useful to look at each of the indirect procedures listed above. This raises interesting and important questions in scientific inference, and as in all such problems, it is vital to get the assumptions straight. The first assumption is that a substance which is a carcinogen in one person is also a potential carcinogen in every other person. We make the further assumption that the carcinogenic potency is about the same in each case, although we know that this is an oversimplification (see below). Since cancer is multicausal, the effect of a given carcinogen may be affected by the previous or subsequent actions of other agents. If one is aware of what the other predisposing factors or agents are, then the assumption is modified to apply only to persons with similar histories of exposure. In the most obvious example, cigarette smokers are found to have higher incidences of many types of malignancies, and to become progressively more susceptible with increasing alcohol consumption (84), which suggests that smoking confers enhanced sensitivity to chemical carcinogenesis.

### Intraspecies Heterogeneity of Responses

Although human genetic heterogeneity is a real and largely

uncontrollable variable in the scaling of carcinogenic potencies from laboratory animals to man, there is evidence to suggest that the differences in susceptibility are not large, except for a small percentage of people who are highly vulnerable. Knudson has suggested that there might exist four groups of persons with qualitatively different susceptibilities to cancer (85). The first group is impervious to environmental effects; the second group has what we consider to be a "normal" susceptibility to carcinogenesis via chemical agents, viruses, and other environmental factors; the third group (which includes persons with XP and Bloom's syndrome) has a genetic susceptibility to environmental carcinogens; and the fourth group has an inherited gene (or developmentally sustained somatic genetic defect) that is equivalent to the first irreversible step along the path toward a particular cancer (85). Knudson points out that there are no distinct lines dividing the first three groups, and that whereas it is now considered likely that most human cancer befalls group two, in fact group two might be a subset of group three, meaning that virtually all susceptibility to environmental carcinogenesis might have a genetic basis (85). We note that unless one understands all environmental variables and controls for them, it is not usually possible to differentiate between a population distribution of genetic susceptibilities and a similar distribution of other factors such as exposure to dietary carcinogens.

That at least two-thirds of all cancers (excluding skin) can be attributed to environmental rather than genetic influences was demonstrated by the pioneering work of Higginson and Oettle, who compared the cancer incidences of the South African Bantu with U.S. blacks (86). More recently, it has been estimated that at least 75 to 80% of cancers are the result of environmental influences, which include smoking, alcohol consumption, hormonal factors controlled by behavior (such as age at first pregnancy), diet, viruses, ionizing radiation, and chemical carcinogens from sundry sources (79,84,87). The fact that, except for rare cases of identifiable individuals at high risk, geographical clusters of cancer attributable to environmental causes seem to be distributed over the exposed population, rather than being concentrated in families or within ethnic groups, suggests that there is a definable upper limit on the range of susceptibilities in the population at large.

A review of human phenotypic variability with respect to the metabolism of several classes of carcinogens and pro-carcinogens showed that in most cases the spread in rates of enzymatic deactivation and activation was within an order of magnitude. For one case in which the distribution was more extreme (three to four orders of magnitude for debrisoquine oxidation), there was also a marked variation among seven rat strains (88). The finding of huge phenotypic variation in some biochemical activities relevant to carcinogenesis is not, however, an indicator of prevalence. In the example noted, 75 to 80% of humans were found to fall within a 10-fold range of debrisoquine oxidation activity. It is also not unreasonable to imagine that, given the overall biochemical heterogeneity of our species, an individual might be more predisposed than average to develop a particular type of cancer and less predisposed to develop another. We expect that the net variability in susceptibilities (within the "normal" population) in each case to be less than the expected error in the scaling of carcinogenic potencies from rodents to humans, which is predicted to be accurate only within an order of magnitude



on average, and is sometimes accurate only to within two orders of magnitude (39,89). Therefore, in most of what follows, exceptions and caveats to our stated assumptions will be ignored. The question of which animal model best simulates the typical human response may be considered when such information is available. Pharmacokinetic studies of some chemicals may also suggest that the animal data extrapolates better to a fraction of the human population which is more (or less) sensitive than average.

### Comparison of Spontaneous Cancer Incidence (Background Tumors)

The use of animal models for human carcinogenesis was originally based, in part, on the general observation that animals in the wild have roughly the same overall rates of malignancy as humans. Use of inbred rodent strains in carcinogen bioassays later became commonplace, and strains have been developed which, intentionally or not, are more susceptible than others to spontaneous and induced neoplasms at particular sites; e.g., liver tumors in male B6C3F<sub>1</sub> mice. In an attempt to maximize the likelihood of detecting a carcinogenic response, ultra-susceptible strains have been used frequently in long-term bioassays. The deliberate use of strains having high, site-specific, spontaneous tumor rates inevitably forces us to reexamine one major rationale for the use of animals as models. A conviction often expressed is that chemical carcinogenesis at some highly sensitive site in an inbred rodent strain does not necessarily predict carcinogenicity at any site in humans. The view held by one of us (G.G.) is that carcinogenic effects in animals can be extrapolated to humans more logically when the animals under test are not subject to unusually high background rates of cancer, whether these increased rates are genetically determined or are artifacts of the experimental design. An alternative view (held by R.W.) is that while it is likely that chemical carcinogenesis in animals with high background tumor rates at particular sites does predict carcinogenicity in other species, the interspecies conversion factor is probably lower for these sites.

Determination of background cancer rates in humans is somewhat of a subjective process. Without a detailed understanding of all the factors which produce tumors, the best that one can usually do is to assume that the lowest rate observed for a given site (in all populations) is the background rate. As discussed above, it has been estimated that at least 75 to 80% of malignancies occurring in people living in the developed nations are attributable to environmental factors, including diet. The percentage of U.S. deaths due to cancer was 22% in 1985, according to the Bureau of Vital Statistics. Based on this figure, the background (nonenvironmental) cancer rate in humans would be less than approximately 0.25 of 22%, or 5.5%. The results of several human autopsy studies suggest that the overall incidence of cancer is higher than 22%; approximately one-third of autopsies revealed cancer, and a surprisingly high percentage of the neoplasms found had been either misdiagnosed or undetected before death (90). Based on this autopsy estimate, the background cancer rate would then be 0.25 of 33%, or 8.3%. However, since human autopsies are not expected to be as thorough as histopathological examinations performed at the end of 2-year rodent bioassays, this estimate is probably still too low.

Table 2 lists control tumor rates (for all sites combined) from the unpublished NTP historical control database for strains commonly used in carcinogenicity bioassays; these rates range from

9% to more than 50%. It has been pointed out that the usual laboratory regime results in overfed, overweight animals with endocrine disturbances and other abnormalities, and that these animals are therefore unsuitable for carcinogenicity testing of chemicals (91,92). It is possible to affect the background tumor rate for laboratory rodents by altering this regime; an easy method is to reduce caloric intake to 75% of what is consumed during *ad libitum* feeding. In one such study of outbred Swiss mice, the incidence of total malignant tumors in otherwise untreated animals was lowered from 11 to 4.4% in males and from 14 to 4.4% in females (91,93). (Results of the Conybeare study are included in Table 2, along with some human cancer rates, for comparison with the NTP control group rates.) In a similar study involving inbred rats, the incidence of malignant mammary tumors was reduced from 25 to 3.8% in females; pituitary adenomas dropped from 24 to 3.8% in males and from 48 to 10% in females (94). These observations suggest that the rates of true spontaneous neoplasms in rodents are similar (within a factor of 2) to those in man when diet is given proper consideration as an environmental variable.

It seems reasonable to hypothesize that the potency of a carcinogen when expressed as a fraction or multiple of background effects will be similar in rodents and humans. We have never seen explicitly stated the argument that comparison of background tumor rates implies that  $K_{hr} \approx 1$ , but it has been frequently implied and used. One of the major inadequacies of this hypothesis lies in its implicit assumption that carcinogenesis via environmental agents is mechanistically similar to spontaneous carcinogenesis. This subject is not yet close to resolution. A less

Table 2. Comparison of overall tumor incidences in laboratory animals and humans.

	No. of animals	% with tumors		
		Malignant	Benign	Total
Control animals for 2-year NTP bioassays <sup>a</sup>				
B6C3F <sub>1</sub> Mice				
Males	1692	42	35	64
Females	1689	45	33	64
F344 Rats				
Males	1596	55	95	98
Females	1643	38	76	88
Osborne-Mendel Rats				
Males	50	26	68	78
Females	50	12	80	88
Sprague-Dawley Rats				
Males	56	8.9	36	39
Females	56	30	68	79
Effect of diet on outbred Swiss mice <sup>b</sup>				
<i>Ad libitum</i> feeding				
Males	160	11		44
Females	160	14		31
75 % of <i>ad libitum</i> feeding				
Males	160	4.4		23
Females	160	4.4		11
U.S. Cancer death rates				
Bureau of Vital Statistics, 1985				
Males		22.5		
Females		21.7		
Estimated spontaneous rates <sup>c</sup>				
Males and females		≅ 5.5		
Estimated upper limit on spontaneous rates <sup>d</sup>				
Males and females		≅ 12		

<sup>a</sup>NTP historical control data.

<sup>b</sup>Conybeare (93), tabulated by Roe (91).

<sup>c</sup>Doll and Peto (84).

<sup>d</sup>From observations on Mormons by Lyon et al. (163).

formidable problem is the discrepancy between animal and human data due to the reporting only of diagnosed human malignancies and not all tumors. Neoplasms that were undiagnosed prior to death and/or did not lead to death are not normally counted.

For the sake of conformity with available data on the occurrence of human neoplasms, one could remove from consideration latent malignancies in rodents that do not lead to death and are detected only during routine biopsy, such as testicular tumors in aged male rats. Alternatively, one may include latent malignancies to maximize statistical significance and then compensate for the elevation of the observed potency so as to derive an interspecies conversion factor applicable to active neoplasms.

### Interspecies Differences in the Tumor Site of Highest Sensitivity

If the background rates of site-specific neoplasms are similar in animals and humans and if malignancies induced by a given chemical agent occur at the same sites in animals and humans, then the interspecies extrapolation would be straightforward. Unfortunately, in reality things are not so simple. Carcinogens often induce neoplastic responses at different sites in different species. Another difficulty in making the animal/human comparison is that most human data on the effects of exposure to chemical carcinogens are for malignancies at only the most sensitive site. This does not imply that other sites are unaffected. For any study of carcinogenesis in humans or animals, choosing the site of the dominant neoplastic response will confer higher statistical sensitivity on the result than would choosing any other site. A bioassay is more sensitive still if an increase in an unusual tumor is found, i.e., one that rarely occurs spontaneously. There is also some evidence to suggest that the interspecies (mouse/rat) correlation is stronger for chemicals which produce rare tumors in rats than for those which produce only common tumors (95).

Initially, scientists often made the assumption that a chemical which causes cancer at a given site in one species is likely, in the absence of information to the contrary, to cause cancer at the same site in other species. Moreover, it seemed to make sense to look for cancer in those organs which manifested lesions in response to toxic insult. For example, when aflatoxin B<sub>1</sub> was shown to cause hepatic necrosis in poultry, pigs, and calves, it was immediately tested for hepatocarcinogenic activity in rats [reviewed by Busby and Wogan (96)]. It is now well known that aflatoxin B<sub>1</sub> is a highly potent rodent carcinogen. Although a clear association between hepatocellular cancer and aflatoxin intake was found in the earliest studies of human populations exposed to high levels of aflatoxin B<sub>1</sub>, interpretation was complicated due to frequent simultaneous infection with hepatitis B virus. But several studies published since 1984 have specifically addressed this question of confounding; it has now been established that aflatoxin B<sub>1</sub> ingestion is a high risk factor for liver cancer, above and beyond hepatitis B infection (97,98). Vinyl chloride is another example of a chemical which was known to be toxic to the human liver before being studied for possible hepatocarcinogenicity. In three rodent species a rare liver tumor (angiosarcoma) has been produced by vinyl chloride inhalation (99,100); angiosarcoma has also been found in humans among vinyl chloride workers (101).

We are interested in the overall increase in cancer in humans due to exposure to a given chemical carcinogen, yet often only the incidence at the most sensitive site is recorded. Is it possible to predict the excess cancer incidence at all sites given the excess incidence at the most sensitive site? If so, then an approximate correction factor may be obtained as follows. It is well known that the principal site of cancer due to smoking is the lung. In the U.S. in 1985, there were 122,700 lung cancer deaths. Of these, approximately 112,240 were due to cigarette smoking (102). But a number of other cancers are also linked to cigarette smoking. The number of U.S. deaths in 1985 due to cancer other than lung cancer was 338,870; of these an estimated 30%, or 101,660, were also attributable to smoking. Thus, for those cancer deaths caused by smoking, the ratio total/lung is  $(101,660 + 112,240)/112,240 = 1.9$ . For vinyl chloride, the excess of malignancies at nonhepatic sites is less than the excess of all liver tumors including angiosarcomas (103,104), and in a recent mortality study of vinyl chloride manufacturing workers, angiosarcomas were found to make up half of the total liver tumors (105).

Crouch and Wilson looked at the ratio of the sum of potencies at all sites to the sum of potencies at the most sensitive site for all chemicals which had been tested by the NCI/NTP as of 1978 and which had been judged to be positive for carcinogenicity; they found that the ratio of the sums was approximately 2 in most cases (unpublished data). For evaluating an interspecies conversion factor, they suggested using only the data for the most sensitive site in both species; statistical validity is improved, and the factors of two tend to cancel (39). A factor of two is small, in any case, compared to other uncertainties inherent in the interspecies correlation.

In practice, the mouse/rat and rodent/human interspecies comparisons do correlate more closely when the tumor site is allowed to vary to accommodate the most sensitive site. In a survey of 58 chemicals, Tomatis *et al.* found a correlation between induction of liver tumors in mouse and induction of tumors at any site in rat and hamster (106). This qualitative correlation was stronger for chemicals that also induce tumors in mice of both sexes at sites in addition to liver. It was later noted that if routes of exposure are similar for animals and humans, then the target sites are more likely to coincide, although more sites in animals are usually affected (107). It was therefore proposed by these authors that induction of any animal tumor should be considered as evidence for possible human carcinogenicity, even if tumors at that particular site have not been shown to occur in humans. A dramatic argument in support of this approach can be found by examining the history of the known carcinogenicity of benzene. This chemical has long been associated with leukemia in humans, but was not observed to produce neoplasms in rats or mice. Now recent studies have shown statistically significant tumor incidences at numerous sites in rodents exposed by ingestion or inhalation (108) and gavage (109). An examination of human data for effects at multiple sites revealed a possible association with tumors at only a few of the sites affected in rodents: multiple myeloma and lymphatic and hematopoietic neoplasia (110).

For some time, the above argument has been the basis of regulatory philosophy, e.g., as described by Anderson *et al.* (111). It is now common practice to evaluate a chemical's carcinogenic potency in animals and humans in terms of the most sensitive site, with the caveat that sites with an extraordinarily high background incidence, such as the liver in the males of some mice strains or the testicles of F344 rats, are sometimes disregarded.

## Should We Include Benign Tumors?

In the regulatory community there is considerable discussion about whether to include benign tumors along with malignant tumors when determining potency in rodents for the purpose of predicting potency in humans. The presence of benign tumors is generally accepted as a highly probable indication of eventual malignancy (97,112,113). If this is the case, then the induction of benign tumors by a chemical should be considered as evidence of that chemical's carcinogenicity. Recently, the relevance of benign neoplasms was evaluated for 143 chemicals tested in NTP bioassays. Only five chemicals produced solely benign neoplasms, and those observed are known to represent transitory or progressive stages in the development of malignancy (114). Qualitatively then, a chemical's ability to induce benign tumors is indicative of its carcinogenic potential. However, in a quantitative statement of rodent carcinogenicity, benign tumors should only be included if they are also included in the definition of the rodent-to-human interspecies conversion factor for carcinogenic potency. If we generally evaluate  $K$  by some procedure without including benign tumors, we might nevertheless wish to include benign tumors for a specific chemical in order to obtain a potency that is statistically significant. We could then use the value of  $K$  derived without benign tumors if we multiply the potency for induction of total tumors (benign and malignant) by the average ratio of malignant to total tumors found in chemical bioassays.

## Human/Rodent Chemical Carcinogenicities

### Qualitative Comparison (Human/Rodent)

In all there are now 50 chemicals, groups of chemicals, or industrial processes for which an IARC Working Group has concluded that there is sufficient evidence of human carcinogenicity (97). For most, the data on human exposures is inadequate to calculate a quantitative risk, and hence there is no opportunity to do a human/animal comparison of their carcinogenic potencies. But we can first ask the simpler question: do these human carcinogens always cause cancer in test animals? Wilbourn et al. (115) evaluated the animal carcinogenicity of 30 chemicals (or groups of chemicals) for which sufficient evidence of human carcinogenicity was reported in *IARC Monographs* volumes 1–41. These are listed in Table 3. (The authors excluded certain industrial exposures which were among the 50 IARC human carcinogens.) Of these 30 human carcinogens, there was sufficient evidence of animal carcinogenicity for 18 and limited evidence for 7; data for 3 were inadequate, and for 2 there were no data at all. The latter 5 agents are arsenic, certain combined chemotherapies (including MOPP [mechlorethamine, vincristine, prednisone, and procarbazine]), conjugated estrogens, smokeless tobacco products, and treosulphan. Note that these authors found that “new data would provide limited evidence” for the animal carcinogenicity of arsenic, referring to the work of Ishinishi et al. (116) and Pershagen et al. (117), which were the first adequate studies in which arsenic was administered via a respiratory route. It was also pointed out that there is sufficient evidence of animal carcinogenicity for some components of

MOPP, namely, nitrogen mustard and procarbazine. Of the remaining 3 agents, the conclusion of Wilbourn et al. was that they “had not been adequately tested in experimental animals and no statement can be made regarding their carcinogenicity in animal models” (115). In the most recent IARC assessment, limited evidence was also found for the carcinogenicity of conjugated estrogens in animals, based on studies published in 1983 and 1984 (97). For arsenic, this IARC report cited a number of studies published since 1981 in support of the conclusion that there now exists limited evidence of its carcinogenicity in experimental animals.

We believe that it makes far more sense to discuss a chemical's carcinogenic potency than the probability or possibility of its carcinogenicity. In cases where the observed increase in neoplasms is not statistically significant, an upper limit to the potency can be derived. The following example illustrates the importance of this approach. Ten years ago, evidence for the carcinogenicity of benzene was considered inadequate or inconclusive. The observation was made that since benzene-induced leukemia in humans occurred at only a very low incidence, studies in animals could not be expected to produce a statistically significant increase in leukemia or any other cancer unless large numbers of animals were exposed (118). This analysis tacitly assumed that the potency in animals was of the same order as that in humans. Up until that time, no adequate study (lifetime exposure, sufficient numbers, high enough doses) of benzene's carcinogenicity in animals had been undertaken. It has been demonstrated since that benzene causes cancer in rodents at rates similar to what was expected from the data in humans (108,109,119,120). Recent experimental evidence prompted Wilbourn et al., in their analysis of the response of animals to human carcinogens, to footnote the “limited” evidence for benzene's carcinogenicity in animals with the statement, “new data would provide sufficient evidence” (115). The absence in 1978 of benzene-induced neoplasia in rodents did not disprove the proposition that the carcinogenic potency of benzene is approximately the same in rodents as it is in humans, although it was sometimes mistakenly thought to do so.

### Quantitative Comparison (Human/Rodent)

In 1977, Matthew Meselson and Sir Richard Doll suggested to one of us (R.W.) the importance of quantitative comparison of the carcinogenic potencies in humans and laboratory animals for chemicals for which human data were available. The results of an initial study were included in a paper by Crouch and Wilson (39). There are several great practical difficulties in undertaking this comparison. The first is that there are a limited number of chemicals that are known to cause cancer in humans. In 1978, there were only approximately 25 substances on this list, including some which are classes of chemicals rather than unique compounds. The second difficulty is that the human data, for the most part, are for uncontrolled, unmeasured, short-term exposures. Consequently, epidemiological studies typically relate cancer to exposure only qualitatively. The third difficulty is that once a chemical has been branded a “human carcinogen,” attempts are made to ban it, and interest in assessing its effects in animals declines. Thus, for many known human carcinogens adequate testing in animals has not been done.

**Table 3. Chemicals for which IARC found sufficient evidence of human carcinogenicity (115).**

Chemical	Evidence in animals
4-Aminobiphenyl	Sufficient
Analgesic mixtures containing phenacetin	Limited
Arsenic and certain arsenic compounds	Limited <sup>a</sup>
Asbestos	Sufficient
Azathioprine	Limited
Benzene	Sufficient <sup>b</sup>
Benzidine	Sufficient
Betel quid containing tobacco (chewing)	Limited
<i>N-N</i> -bis(2-chloroethyl)-2-naphthylamine (Chlornaphazine)	Limited
<i>bis</i> (Chloromethyl)ether and technical chloromethyl methyl ether	Sufficient
1,4-Butanediol dimethanesulphonate (Myleran)	Limited
Certain combined chemotherapy regimes for lymphomas including MOPP <sup>c</sup>	No data
Chlorambucil	Sufficient
Chromium and certain chromium compounds	Sufficient
Coal tars	Sufficient
Coal-tar pitches	Sufficient
Conjugated estrogens	Limited <sup>a</sup>
Cyclophosphamide	Sufficient
Diethylstilbestrol	Sufficient
Melphalan	Sufficient
Methoxsalen with UV A (PUVA)	Sufficient
Mineral oils (containing various additives and impurities)	Sufficient
Mustard gas	Limited
2-Naphthylamine	Sufficient
Shale oils	Sufficient
Smokeless tobacco products	Inadequate
Soots	Sufficient
Tobacco Smoke	Sufficient
Treosulphan	No data
Vinyl Chloride	Sufficient

<sup>a</sup>Listed as "inadequate" in *IARC Monographs* series prior to 1987; see text for details.

<sup>b</sup>Listed as "limited" in *IARC Monographs* series prior to 1987; see text for details.

<sup>c</sup>MOPP, mechlorethamine, vincristine, prednisone, and procarbazine.

We show in Figures 1A and 1B a comparison of carcinogenic potency in humans to that in rats and mice, respectively, taken from Crouch and Wilson (39). In most cases these authors had to estimate human exposures by reanalyzing the data from the original epidemiological literature. Consequently, the results contain considerable uncertainty. The error bars delimit one standard deviation. Where a bar has an arrow at the end, one standard deviation does not define a lower limit; the error in such cases encompasses zero potency. The lines  $\log(\beta_h) = \log(\beta_r)$  and  $\log(\beta_h) = \log(\beta_m)$  (i.e.,  $K_{hr} = K_{hm} = 1$ ) do not pass through the error bars on all the points. Nonetheless, the results are in rough agreement with the proposition that the interspecies factor for rat to human ( $K_{hr}$ ) and for mouse to human ( $K_{hm}$ ) are each about equal to 1, with most deviations falling within an order of magnitude. This range is small compared with the range of carcinogenic potencies in rodents, which vary over five orders of magnitude. We note that the U.S. Environmental Protection Agency (EPA) uses  $K_{hr} = 5.9$  and  $K_{hm} = 13$  in its risk assessment procedures (Table 1); by inspection it can be seen that these EPA values for the interspecies conversion factors are more likely to overestimate the risk. Although the simplest proposition (that  $K_{hr} = K_{hm} = 1$ ) is almost certainly not precisely true, the proposition that the values of  $K_{hr}$  and  $K_{hm}$  are lognormally distributed

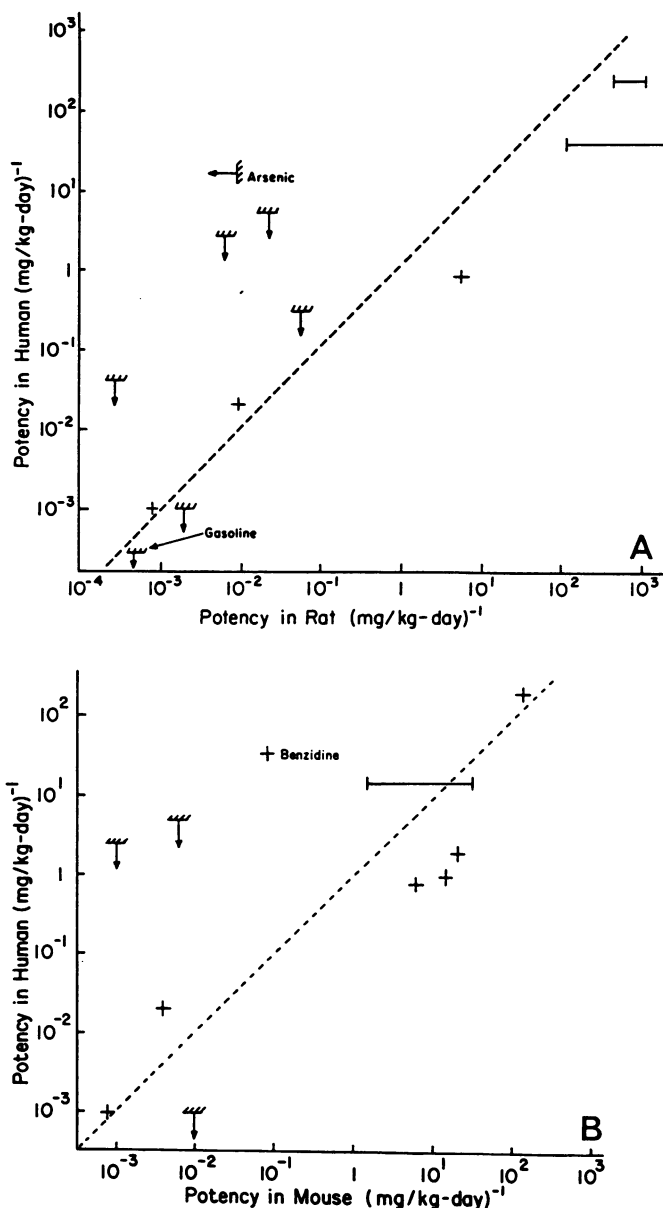


FIGURE 1. Carcinogenic potency in humans versus rats (A) and mice (B). From Crouch and Wilson (39). The point for gasoline, calculated by us, has been added to 1B. (See text for details.) Otherwise, the data reflect information available in 1979.

about 1 is consistent with the data. Letting  $x$  represent  $\ln(K_{hr})$  and  $\ln(K_{hm})$ , the probability of finding a value  $x$  is given by

$$P(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-x^2}{2\sigma^2}\right) dx \quad (2)$$

where the standard deviation  $\sigma$  is to be determined.

Allen et al. performed a much more extensive survey of human carcinogenic potencies and made comparisons with rodent potencies (89). As noted above, probably the most difficult aspect of such a study is the evaluation of human exposures. The best estimates of the human  $TD_{25}$  values (with their associated uncertainties) for the 20 chemicals for which sufficient data were found is reproduced from Allen et al. (89) in Figure 2. Unfor-

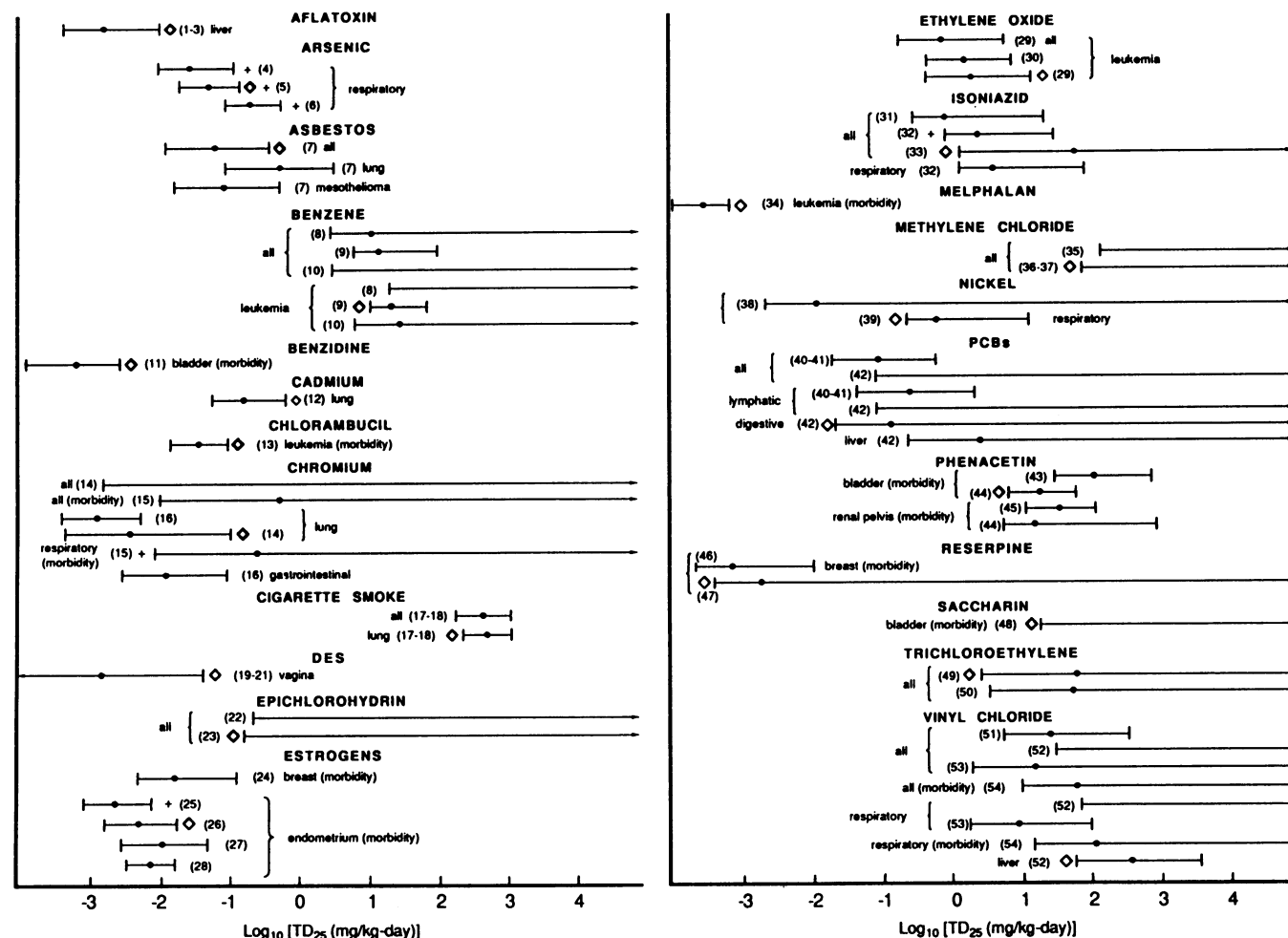


FIGURE 2. Best estimates and upper and lower bounds for  $TD_{25}$  values from each human study. Numbers in parentheses are citations which may be found in the original publication. ( $\diamond$ ) Data selected to represent the chemical when comparisons with bioassay-based estimates are made; (+) poor fit of linear dose-response model that is not due to lack of response. From Allen et al. (89).

tunately, the detailed calculations upon which these estimates were based are unpublished, so that we are unable to judge their reliability. In Figure 3 we show their base case plot of human versus rodent  $TD_{25}$  estimates. We note here that they have not derived a best fit line in the ordinary way. Their statistical significance statement is for how well the order of the potencies in rodents correlates with the order of the potencies in humans. However, by inspection of Figure 3 we deduce that the numerical values of the potencies also correlate reasonably well. The solid line corresponds to  $K_{hr} = K_{hm} = 1$ ; the dotted line,  $K_{hr} = (W_h/W_r)^{1/5} = 5.9$ ; the dashed line,  $K_{hm} = (W_h/W_m)^{1/5} = 13$ . It can be readily seen that  $K_{hr} \approx 1$  and  $K_{hm} \approx 1$  are more likely propositions than  $K_{hr} = 5.9$  or  $K_{hm} = 13$ . The relationship for  $P(x)$  in Eq. [2] fits the data reasonably well with  $\sigma \approx \ln(5)$ .

While making this statement, we emphasize that neither Crouch and Wilson (39) nor Allen et al. (89) have proven that  $K_{hr} = 1$ . These authors merely tested a proposition and found it consistent with the available data. This work and related work on the correlation between rat and mouse carcinogenic potencies have been criticized by Freedman and Zeisel, who highlight the

limitations of the available data and insist that the proposition that  $K$  is distributed around the value of 1 is not proven (121). But in fact, the proposition has not been disproven for any chemical for which adequate data are available for both human and rodent exposures. However, for most chemicals the above proposition remains untested. There may be chemicals for which the rodent bioassay is a poor predictor of human carcinogenicity, with potencies in humans (if they were measured) such that  $K_{hr} \gg 1$  or  $K_{hr} \ll 1$ . We know of no such chemicals, but this does not mean that they do not exist.

## Negative and Inadequate Epidemiology

When we attempt to compare animal and human carcinogenic potencies, and in particular, when we try to understand the causes of any discrepancies, we are impressed by the importance of "negative epidemiology," i.e., studies where no statistically significant effect was observed. The absence of a significant effect does not, of course, prove that the chemical does not cause cancer; rather, any increased incidence is too small to be deter-

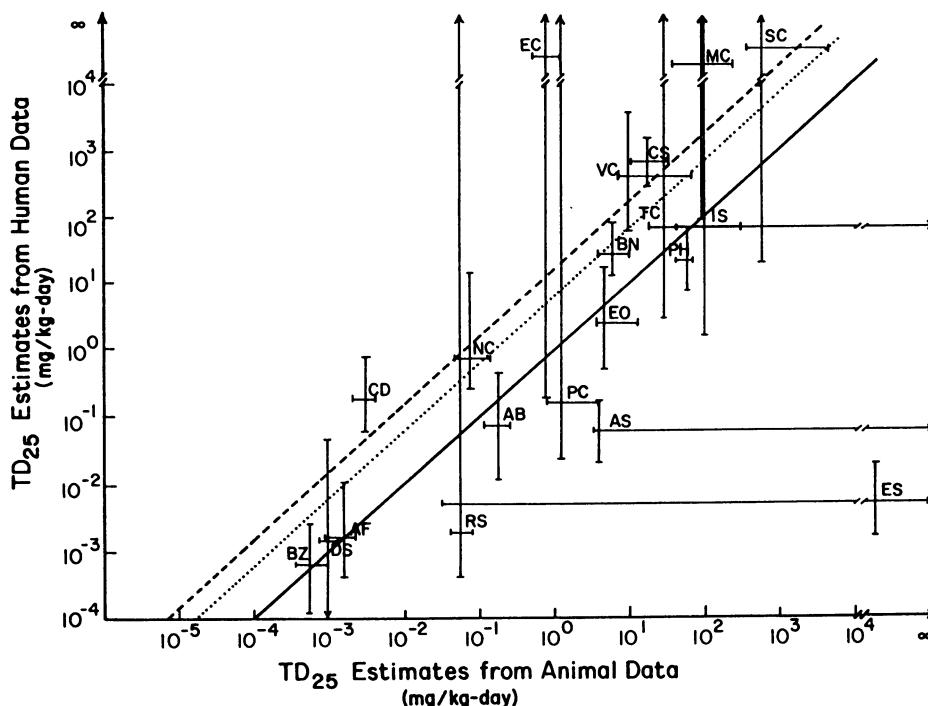


FIGURE 3. Human  $TD_{25}$  estimates versus animal  $TD_{25}$  estimates. From Allen et al. (89).

mined unambiguously under the given exposure conditions. It becomes useful to express the result as an upper limit to the number of cancers that could have been caused, and then, if the exposure level is known, to derive an upper limit to the carcinogenic potency, and therefore an upper limit to  $K_{hr}$ , the lower limit being zero. Crouch and Wilson introduced this concept in their comparisons (39), and it was also used by Allen et al. (89). We believe that there are many more chemicals to which it might be applied. For example, contact with chlordane and heptachlor did not increase the age-corrected lung cancer mortality in a group of termite control workers, who were more likely to incur higher exposure to these agents than were any other segment of the population (122). Although this study did not correct for smoking, an estimated upper limit on potency might still be extracted.

We shall illustrate next how a probable limit on human carcinogenic potency can be derived from inadequate epidemiological results. Enterline and Viren have reviewed the evidence for the association between kidney cancer and exposure to gasoline fumes (123). They concluded that the bulk of evidence from cohort studies of petroleum-refining or distribution personnel indicates a small excess of kidney cancers among older workers exposed for long periods. Let us do a back-of-the-envelope calculation of the upper limit to the potency of gasoline fumes for induction of renal cancer, based on data cited by these authors. Exposure levels were drawn from a study of the ambient concentration of hydrocarbons at gasoline marketing terminals; the mean value of the concentration as determined by personal samplers (worn  $\geq 7$  hr) was found to be 5.4 ppm (124). The highest renal cancer death rate reported for this industry was 0.15% for workers who had been exposed for 20 years or longer;

this excess was statistically significant (125). The expected mortality rate from renal cancer for the whole U.S. population was 0.071%. Assuming an average molecular weight of 80 daltons for the volatile constituents, then for a 70 kg man working 8 hr with inhalation volume 20.8 L/min, the daily dose would be 2.8 mg/kg. Making the conservative assumption of a linear dose response between 0 and 2.8 mg/kg-day, then the potency  $\beta \approx (0.0015 - 0.00071)/(2.8 \text{ mg/kg-day}) = 2.8 \times 10^{-4} (\text{mg/kg-day})^{-1}$ . (For simplicity we neglect correcting for length of working life, which would result in a lower predicted potency.)

For calculation of a corresponding carcinogenic potency in rodents, we turn to a lifetime study in which rats and mice were exposed to volatilized gasoline for 6 hr/day (126). A dose-related increase in renal cancer was found in male rats; interpretation was simplified by the fact that the strain of rats used, F344, has a very low incidence of spontaneous kidney tumors. In dose groups of 100 animals each, at dose levels of 0, 67, 292, and 2056 ppm there were, respectively, 0, 1, 5, and 7 malignant neoplasms. Because the gasoline was vaporized to completion by heating, this study is not a perfect model for the human exposures. Adjusting the molecular weight estimation to 128 daltons to reflect the mostly  $C_9$  composition of commercial gasoline, and assuming a weight of 0.5 kg and an inhalation volume of 0.10 L/min, the daily dose was 120 mg/kg. Since the response is essentially linear in the range of 0 to 292 ppm, we calculate the potency as  $\beta \approx (0.05 - 0.0)/(120 \text{ mg/kg-day}) = 4.2 \times 10^{-4} (\text{mg/kg-day})^{-1}$ . (We assume here that there was one neoplasm per rat. If this was not the case, then the incidence rate would be lower, as would the potency estimate.) In Figure 1A, we have added the data point just calculated to a figure published originally by Crouch and Wilson (39). With very little effort we have been able to increase the body



of quantitative knowledge pertaining both to the carcinogenic potency of gasoline in humans and to the interspecies correlation factor.

There is an obvious tendency to concentrate on those chemicals where a definite effect in humans has been found, and indeed upon these the correlation principally depends. However, if a chemical is found to be carcinogenic in animals and is not found to be carcinogenic in man, it is of obvious interest to discover the cause of this negative outcome. Is it because the chemical is an exception to the correlation; people were not exposed at sufficiently high levels or for sufficiently long durations or because the available epidemiological data are inadequate?

## Chemicals Not Known to Be Carcinogenic to Humans

Ennever et al. discuss 29 chemicals for which at least one epidemiological study had found no evidence of human carcinogenicity. They looked at the quantitative (+/-) evidence for rodent carcinogenicity for 20 of these chemicals. They reported that only one, methotrexate, was negative in rodents, and that the remaining 19 were positive. They concluded that the specificity of rodent bioassays for predicting human noncarcinogens is very low (127).

Whereas Ennever et al. looked for qualitative agreement between the rodent carcinogenicity bioassay and human data, we were able to ask more meaningful, quantitative questions about the agreement by examining the actual TD<sub>50</sub> values (128). Starting with the rodent TD<sub>50</sub> at the most sensitive site from the Carcinogenic Potency Database (CPDB) (129-131), we derived a predicted human incidence for the degree of exposure and duration of follow-up corresponding to the most comprehensive epidemiological study available and then compared the predicted incidence with the observed incidence. If a chemical produced no statistically significant increase in cancer at any site in the exposed population, consistency with rodent results is inferred if the minimum rodent TD<sub>50</sub> is sufficiently high that no attributable cases would have been expected under the actual conditions of human exposure and follow-up. For 18 of the 22 chemicals examined, the human evidence is consistent with the predictions based on the rodent bioassay results. For two chemicals, dichlorobenzidine and ethylene thiourea, there is not enough epidemiological information to make a useful comparison with rodent bioassay data. The two chemicals for which the human evidence is inconsistent with the predictions are actinomycin D and vinylidene chloride. But for actinomycin D, the conditions of the rodent bioassay were inappropriate for the comparison, and for vinylidene chloride the human exposure dose was uncertain; for either chemical future studies in humans might yet demonstrate consistency with the rodent results (128).

## Rat/Mouse Chemical Carcinogenicities

Of the chemicals that have been tested in animals, there are few for which reliable information as to human carcinogenicity exists. Many more chemicals have been tested in two rodent species, and hence there is a greater opportunity for performing either qualitative or quantitative interspecies comparisons of carcinogenicity. Most assays are performed in rodents, primarily because of their relatively low cost, ease of maintenance, and

short lifespan. Their use may be fortuitously appropriate in some cases. In one study of the efficacy of animal tests in the qualitative (yes/no) prediction of human toxicity for 20 chemicals, production of toxic lesions was found to be twice as predictive when the test species was the rat or the mouse as it was when dogs were tested, and limited data showed approximately the same correlation for monkeys as for these rodent species (132). Also on the basis of comparative toxicology, it has been suggested recently that the results in rabbits might extrapolate more accurately to humans than do results in rats (14). Interspecies comparisons are facilitated when all the animal tests are performed uniformly. The NCI/NTP tests of rats, mice, and hamsters conform to a standard protocol and therefore lend themselves to such comparisons (133,134).

## Qualitative Comparison (Rat/Mouse)

The concordance of positive/negative results for carcinogenicity between rodent species has been examined by several authors. The fraction concordant is defined as the number of chemicals positive (for at least one site) in both species plus the number of chemicals negative in both species, divided by the total number of chemicals tested. The NCI/NTP conclusions as to carcinogenicity for 266 chemicals adequately tested in both species were tabulated by Haseman and co-workers. With equivocal results excluded, 167/212 (76%) were concordant (6,7). Details of the correlations are reproduced from Haseman and Huff in Table 4. These results were similar to those of Purchase, who reported 82% concordance for a similar number of experiments (135), many of which were included in the later study. Byrd et al. reanalyzed the raw data of the NCI/NTP data base. They found that 76% of the results were concordant when weakly significant ( $p < 0.025$ ) dose-response trends were counted as positive. They also looked at how well positivity/negativity for benign or malignant mouse liver tumors predicts the presence/absence of a carcinogenic response in the rat at any site: the overall concordance was 155/290 (53%) for benign liver adenoma and 165/290 (57%) for malignant liver carcinoma (8). When all the available bioassay data tabulated in the CPDB (not limited to the NCI/NTP subset) was evaluated with respect to the same correlation (for malignant and benign tumors combined), Gold et al. (136) found a concordance of 74/117 (63%). With chlorinated chemicals (defined as compounds consisting only of C, H, Cl, and optionally O) excluded, the concordance is improved: 60/86 (74%).

Gold et al. also examined the overall positive/negative agreement for carcinogenicity between rats and mice for the

Table 4. Correlations in tumor response for 266 NCI/NTP carcinogenicity studies.\*

Comparison	+/+	+/-	-/-	-/-	Total	% Concordant
Male rats/female rats	74	25	12	181	292	87.3
Males rats/male mice	46	43	36	145	270	70.7
Male rats/female mice	59	33	36	145	273	74.7
Female rats/male mice	46	32	37	156	271	74.5
Female rats/female mice	57	23	39	156	275	77.5
Male mice/female mice	78	10	23	177	288	88.5
Rats/mice (overall)	67	32	36	131	266	74.4

\*Data reproduced from Haseman and Huff (6). Equivocal outcomes were considered negative.

CPDB chemicals (136). Of the 392 chemicals tested in both species, 226 were positive in at least one sex/species group; 76% of the rat carcinogens were positive in the mouse, and 75% of the mouse carcinogens were positive in the rat. The overall concordance was 76%. Their criteria for positivity were the same as those of Haseman and Huff (6). Experiments were classified as positive only if this was the published opinion of the author(s) of the original study; all other experiments were classified as negative. No quantitative dose-response or potency information was utilized. The predictivity of the ten most common target sites was also examined. Most sites were found to be good predictors of carcinogenicity at some site in the other species. The mouse liver and the rat urinary bladder/urethra were the least accurate predictors of a positive carcinogenic response in the other rodent species, and the predictivity for these sites is better for chemicals that produce tumors at another site as well. Chlorinated compounds were significantly less predictive than other mouse liver carcinogens: 45% (14/31) of those positive in mouse liver are positive in the rat, compared to 70% (60/86) of other compounds.

The concordance between rats and mice for a "random" sample of 25 chemicals drawn from the first 192 NCI bioassays was assessed independently by three groups of statisticians as part of a symposium on statistical problems; the results were introduced and summarized by Young (137,138). The NCI had previously declared 46% of the 25 chemicals to be positive for carcinogenicity in at least one sex/species group, with 76% agreement between rats and mice. All three groups of statisticians evaluated the studies on a positive/negative basis for carcinogenicity; potency information was discarded.

The first analysis, the decision-tree approach of Sanathanan et al., is the most satisfying on the basis of current biological understanding and what logic tells us to expect for the minimal manifestations of carcinogenicity. They found an overall agreement of 75% between rats and mice for tumors at any site (139).

An even higher rate of agreement (83%) was achieved in the second analysis, by Louis, which combined data over both sexes and which included only malignant tumors (140). Several noteworthy approaches to combining bioassay data were included, although the author's Bayesian combination of potency information from all chemicals is puzzling to us, since no adjustments were made on the basis of chemical structure or pharmacokinetic parameters.

The third analysis, by Bickis and Krewski, which is actually four separate sub-analyses, was apparently conceived as a means of pointing out to others what they ought not do (141). Three of the authors' subanalyses rely on aggregate tumor data; in the remaining one, their Decision Rule III, data for specific lesions is tabulated (which has been shown by several other studies cited earlier to be the better approach). When mock historical control incidences were included, using Decision Rule III resulted in 96% (24/25) of the chemicals being found positive for carcinogenicity in both species; the remaining chemical was positive in rats and inconclusive in mice. Therefore the agreement between rats and mice was 100% (24/24). When historical incidences were not considered, then with this decision rule only 14 of the studies gave adequate results, and only 6 of these (43%) were concordant in rats and mice. The authors make some interesting inferences about false negatives, false positives, and the pitfalls of using historical control data; none of the four sub-

analyses would be selected by anyone who, upon rational consideration, is seriously intent upon getting the most reliable information out of the rodent bioassay. We emphasize this judgment of ours because the figure of 96% positivity has misleadingly been included in the abstract of an introduction to these papers (137), referring to the "most liberal" decision rule of Bickis and Krewski (141), which those authors themselves surely would never recommend. Their four decision rules were chosen as an exercise in illustrating the inadequacies of various simplifications, rather than as complete methods in themselves. Thus, Young (137) has manufactured a huge inconsistency in the analysis of bioassay data, especially with regard to the decision of positivity, where none really exists.

Haseman has recently evaluated the entire analytical exercise in detail (142). He points out that only very limited information was available to the statistical analysts: summarized site-specific tumor incidences, survival rates, and sex. All other information was withheld, including individual animal data, time-to-tumor, the identity of the chemical compounds, dose levels, tumor descriptions, and historical control data. Thus, the analysts were forced to rely upon nonstandard methods, which were less refined than what are normally used (142). We note also that the results would have been more meaningful if the bioassay data for a much larger group of chemicals, say the entire NCI data set, had been subject to differential statistical analysis.

## Quantitative Comparison (Rat/Mouse)

Many toxicologists have only considered a chemical to be a potential human carcinogen if it has caused a significant number of tumors in two different animal species—usually rats and mice. It is this set of chemicals for which one performs quantitative studies of correlations between carcinogenic potencies in the two species. This leaves open the question of the meaning of the results of a bioassay in only one species, or bioassays in two species where a statistically significant number of tumors was found in only one of the species. Is the chemical a carcinogen in one species and a noncarcinogen (i.e., has zero carcinogenic potency) in the other? Or is the failure to find evidence of carcinogenicity in the second species due to the limited sensitivity of the bioassay?

A statistical study cannot answer this question directly. But if we reformulate the problem in mathematical terms then it can be more readily addressed. We begin by positing a model in which the carcinogenic potency in species *b* can always be derived from that in species *a* to within a certain accuracy, according to the formula

$$\ln\beta_b = \ln\beta_a + \ln K_{ba} + \epsilon \quad (3)$$

where  $\epsilon$  is a random error variable given by

$$P(\epsilon)d\epsilon = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right) d\epsilon \quad (4)$$

Then we ask, in those cases where a statistically significant increase in tumors is found only in one species, does this contradict the model?

Crouch and Wilson took a qualitative argument, that the induction of tumors at one site in one species is an indicator of tumorigenicity at different sites in another species and made it quantitative (39). They calculated maximum likelihood esti-

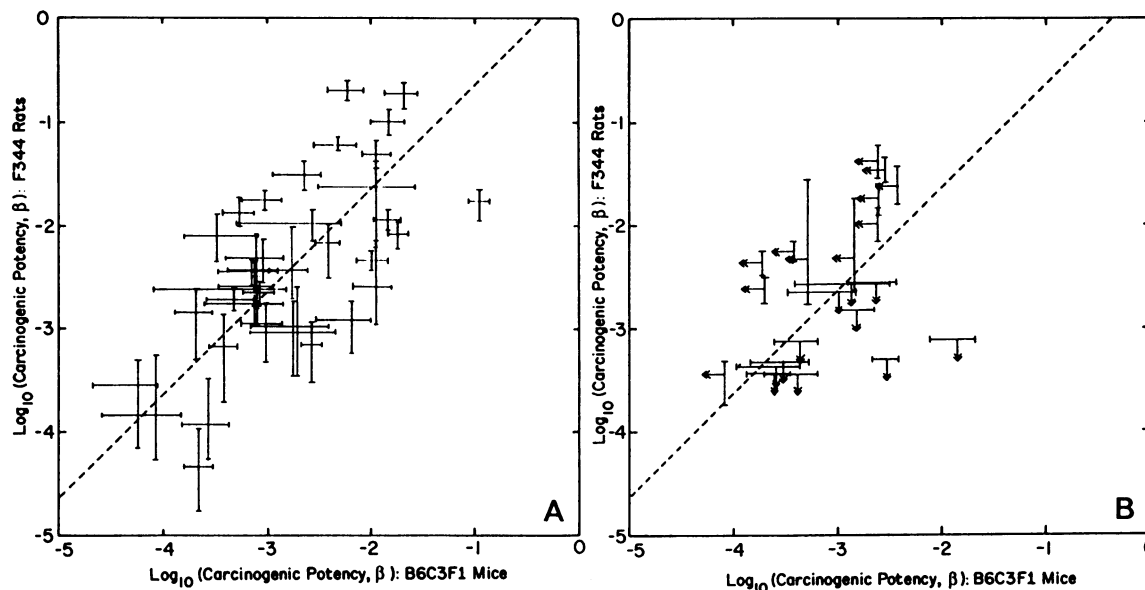


FIGURE 4. Carcinogenic potency in B6C3F<sub>1</sub> mice versus F344 rats. (A) Dashed line is best fit of data shown. (B) Cases that are not in evident agreement with the dashed line of 4A. From Crouch (37).

mates (MLE) of  $\beta$  from NCI/NTP bioassay data, took the geometric mean of the potencies at the most sensitive site in males and females, and compared these sex-averaged potencies for rats and mice. They found that the potencies were in better agreement if the comparison was based on the most sensitive site in each species than if the site was fixed. Crouch evaluated the correlation analytically using Eq. (3) and found the best fit solution for the interspecies conversion factor  $K_m$  between rats and mice (37). For 42% (78/187) of the chemicals there was a statistically significant ( $p \leq 0.025$ ) tumorigenic response in one or both species. Of these, 47% (37/78) were significant in both species; the correlation is shown in Figure 4A. For 19 of the 39 chemicals which produced a significant response in only one species, the results were in self-evident agreement with Figure 4A; results for the remaining 22 are shown in Figure 4B. As before, the error bars encompass zero when the data are not statistically significant. Although some of the error bars do not touch the best fit line (obtained from the experiments which produced a significant response in both species), the deviation is not large. All the data are consistent with the correlation given by Eq. (3), with the error term defined in Eq. (4), with  $\sigma$  chosen to fit the data. In each case the error  $\epsilon$  in the correlation equation is greater than the statistical uncertainty of the individual points.

Crouch showed that there are few cases, if any, where one can say definitively that a chemical is carcinogenic in one rodent species and not in the other, based on the NCI/NTP bioassay results (37). More precisely, he found that the carcinogenic potency in one species is rarely more than 100 times less than the potency in the other. Metzger et al. broadened the scope of the analysis by looking at bioassay data other than that obtained by the NCI/NTP program; approximately the same interspecies (mouse/rat) correlation of minimum  $TD_{50}$  values (maximum potencies) was found in both NCI/NTP and non-NCI/NTP datasets (143). Rieth and Starr also found strong interspecies (mouse/rat) correlations for a) maximum carcinogenic potencies

( $1/TD_{50}$  values) for chemicals for which the authors' opinion was positive for carcinogenicity, and b) upper bound potencies (based on lower-bound estimates of  $TD_{50}$ ) for chemicals for which the authors' opinion was negative with respect to carcinogenicity (144). They found weaker interspecies correlations for the cases where the chemical was positive in mice but not rats, or positive in rats but not mice. It seems to us that these results are in accord with those of Crouch (37) and that the correlations described are consistent with the hypothesis that the interspecies comparison of potencies is meaningful even for chemicals which produce strong evidence of carcinogenicity in only one of the two rodent species. However, Rieth and Starr have a different interpretation. They point to the interspecies correlation of MTDs as a source of bias (see below).

The findings of Crouch and Wilson (39) were confirmed by Gaylor and Chen (145); some of their results are reproduced in Table 5. They compared the minimum  $TD_{50}$  values in rats, mice, and hamsters for all the chemicals in the original CPDB (129) as a function of route of administration and tumor site. They reported that the geometric mean of the ratio for mice/rats is 2.2 for diet and 1.13 for gavage. When the tumor site for both species was the liver, then the geometric mean (mice/rats) is 1.48 for diet. Inhalation gave the poorest interspecies correlation. When the route was diet, the ratio of minimum  $TD_{50}$  values (mice/rats) was distributed such that 73% (138/190) were between 0.1 and 10, 1.6% (3/190) were greater than 100, and 0.53% (1/190) were less than 0.01. Similar ratios were found for the hamster/rat and hamster/mouse comparisons, although data was limited. The authors concluded that the variation in the minimum  $TD_{50}$  values across the three rodent species is generally within a factor of 100 over a wide range of chemical compounds.

Chen and Gaylor (146) used the Crump linearized multistage model to find the upper confidence limit on potency (for the most sensitive site) at the lowest experimental dose for 38 NCI/NTP carcinogens; from the confidence limit they calculated the

Table 5. Ratios of minimum TD<sub>50</sub> values (mg/kg-day), from Gaylor and Chen (145).

Species	Route	Tumor site	No. of compounds	Ratio <sup>a</sup>	SD (log <sub>e</sub> )
Mouse : rat	Diet	All	190	2.20	1.90
Mouse : rat	Diet	TBA <sup>b</sup>	33	2.22	1.58
Mouse : rat	Diet	Liver	16	1.48	1.61
Mouse : rat	Gavage	All	32	1.32	2.39
Mouse : rat	Gavage	TBA <sup>b</sup>	11	4.22	1.61
Mouse : rat	Water	All	10	0.69	2.42
Mouse : rat	IP	All	16	1.41	1.24
Mouse : rat	IP	TBA <sup>b</sup>	12	1.03	1.14
Mouse : rat	Inhalation	All	7	11.2	3.28
Hamster : rat	Diet	All	10	13.7	2.40
Hamster : mouse	Water	All	7	8.49	1.57

<sup>a</sup>Ratio of geometric means.

<sup>b</sup>TBA, tumor-bearing animals, any site.

estimated "virtually safe dose" (VSD), defined as that dose which produces a cancer risk of  $10^{-6}$ . For 69 same-sex comparisons between B6C3F<sub>1</sub> hybrid mice and rats of three strains, the VSD ratios (rats/mice) varied from 0.020 to 49, with geometric mean 1.27 and standard deviation  $\ln(6.23) = 1.83$ . The ratios of the minimum VSD values without regard to sex varied from 0.021 to 26, with geometric mean 1.37 and standard deviation  $\ln(5.99) = 1.79$ . The VSD ratio  $r$  was distributed such that  $0.01 \leq r < 0.1$  for 12% (8/69), and  $10 < r \leq 100$  for the remainder (8/69).

## Biases

Almost all experiments and correlations have biases, and it is important to discuss them. One obvious potential bias in the interspecies comparison of carcinogenic potency is that we normally look only at chemicals that have been found to produce a statistically significant excess of tumors in both species. This leaves open the question: Are there chemicals which are carcinogenic in only one of two species in which they were tested? If such chemicals exist, then the correlation is obviously limited. We addressed the question of the limit of sensitivity of the bioassay by plotting the carcinogenic potencies in mice versus those in rats, along with the associated statistical uncertainty, in Figure 4. The error bars show that in every case there is an upper limit to potency, but the lower limit sometimes encompasses zero (Figure 4B). From the magnitude of the uncertainties it was concluded that for these examples, the occurrence of statistically significant ( $p \leq 0.025$ ) evidence of carcinogenicity in only one species did not provide evidence of exceptions to the correlation. However, Figure 4 only indicates consistency with the correlation; it does not prove that the correlation is in fact followed for every chemical. Some chemicals could be exceptions to the rule, but it is not possible to tell.

There has been no systematic attempt, to our knowledge, to study the biases in the qualitative (yes/no) interspecies concordance studies of the type conducted by Haseman and Huff (6), Gold et al. (136), and others, in which carcinogenicity is scored on a positive/negative basis. Yet simple arguments suggest that the biases will be greater than those for the quantitative comparisons. As evidenced in the above paragraph, one virtue of the quantitative procedure is that it permits interspecies comparison of chemicals positive in one species and nominally negative in another (Fig. 4B). The qualitative interspecies concordance

studies are, of necessity, silent about these cases.

As we have already noted, the experimental design of the typical bioassay is such that, for most chemicals, the carcinogenic potency at the most sensitive site is just above the limit of sensitivity. This means that the interspecies concordance of carcinogenicity/noncarcinogenicity is very sensitive to the criteria for positivity in the individual bioassays. Recall the four decision rules set down by Bickis and Krewski (141), the most conservative of which leads to a high percentage of false negatives (the assignment of non-carcinogenicity to chemicals that would logically have been labeled as carcinogens under any set of reasonable positive/negative criteria), and the most liberal of which leads to the assignment of positivity in 100% of the cases, a large number of which are false positives (chemicals of very low or zero potency that would logically have been labeled as noncarcinogens under any set of reasonable positive/negative criteria). If statistical decision rules for carcinogenicity are not tempered by judicious use of biological information, the outcome is far less convincing than when all data, however untidy, is allowed to play a role. On the other hand, if the details of the requirements for positivity/negativity are not decided in advance, there is enormous potential for bias (147). A further bias in positive/negative decisions enters when the experimenter has a desire to prove that a chemical is carcinogenic. The more bioassays that are carried out, the greater becomes the probability that one will produce a significantly positive outcome, merely due to chance. It is our contention that biases in yes/no concordance studies can be large in either direction, and that the best way of studying these biases is by looking at the quantitative potency relationships.

Biases can also appear in attempts to find qualitative correlations between genotoxicity and carcinogenicity. In a study of 73 chemicals tested in recent NCI/NTP rodent bioassays, Tennant et al. found that there was no complementarity between four commonly used *in vitro* tests for genotoxicity (23). The Salmonella mutagenesis test had the highest specificity (negative response to nominal noncarcinogens) and the lowest sensitivity (positive response to carcinogens). They concluded that no battery of tests constructed from these four tests offered an improvement over the Salmonella assay. Sensitivity could be improved, but then specificity was sacrificed. Furthermore, the three most potent carcinogens examined were not positive in any of the *in vitro* tests. As others have realized, if a chemical is negative in a series of short-term tests having distinct genotoxic end points, further testing sometimes amounts to the experimenters' relentless pursuit of just one positive genotoxic response. If enough tests are attempted, the laws of probability, applied to the tests' false-positive rate, make it likely that a positive result will be found in some test.

That there is a good correlation between the carcinogenic potency at the most sensitive site in rats with that in mice is now firmly established, as has been elaborated in this review. But the basis and relevance of this correlation has been the subject of heated discussion. In particular, a paper by Bernstein et al. (41) has been misinterpreted frequently as providing an argument against the validity of the interspecies correlation. Here we shall endeavor to explain the results of that paper (quoting the authors directly) and the argument it engendered. Most of the following points have been made previously by Zeise et al. (148-150) and Crouch et al. (42). We hope that our expository efforts here will

be successful at clarifying once and for all our position on what we believe has been proven and what has not.

Bernstein et al. found that the best estimate of  $b$  of the carcinogenic potency ( $\beta$ ) of a chemical tested in a rodent bioassay can be approximated as a simple function of the maximum dose tested (MaxD). This may be written as

$$b \approx \frac{\ln(q_0/q)}{\text{MaxD}} \quad (5)$$

where  $q_0$  is the fraction of tumor-free animals in the zero-dose group, and  $q$  is the fraction of tumor-free animals in the MaxD group. If for rats and mice inclusive the possible range of values of  $\ln(q_0/q)$  is much smaller than the range of MaxD values, and if there is an interspecies correlation between the MaxD values, then Eq. (5) implies that there would have to be an incidental interspecies correlation between the carcinogenic potencies as well.

Bernstein et al. did indeed find a good same-sex correlation for MaxD values between rats and mice in 186 NCI bioassays (41). As stated above, the MaxD is usually just the MTD or is close to it in value, while the MTD is a measure of a chemical's chronic toxicity. The existence of an interspecies (rat/mouse) correlation for acute toxicities ( $\text{LD}_{50}$  values) had already been described (148,151).

Bernstein et al. then attempted to quantify the expected variation in  $\ln(q_0/q)$  for an ideal two-group experiment in order to predict the possible outcomes of the carcinogenic potency (41). In their hypothetical scheme, there are 50 animals in each of 2 dose groups; the control group receives dose = 0, the treatment group receives dose =  $d$ . They further assumed a 10% incidence of tumors in the control group ( $q_0 = 0.9$ ) and reasoned that there must be at least a 20% incidence of tumors in the treatment group ( $q \leq 0.8$ ) in order that there be a statistically significant increase in tumors. Initially, they allowed for the possibility of 100% tumors ( $q = 0$ ). With these entirely reasonable simplifications, Bernstein et al. concluded that "the range of possible statistically significant estimates of  $[\beta]$  spans  $0.118/d$  to infinity." This means that the range of possible values of  $\ln(q_0/q)$  is not small compared to the range of MaxD values.

But we know that only for a few chemicals (of which ethylene dibromide is an example) does the fraction with tumors approach 100% at any site. This is a fact of nature; it has nothing to do with experimental design. There is no reason why there should not exist a carcinogen with low toxicity (high MTD), like saccharin, and high carcinogenic potency, like TCDD. Such carcinogens may exist, but none has been found! We emphasize that a carcinogen having these properties could readily be detected in a standard rodent bioassay; at the MTD it should yield tumors in 100% of the animals under test at the most sensitive site.

The opposite situation, in which a chemical has high toxicity relative to its carcinogenic potency, may very well exist but would not be detected in a standard bioassay; at the MTD there would not be a significant increase in tumors. Of chemicals tested in animal bioassays, approximately half are not found to be carcinogenic (6-8,31,136). These produce either no increase or some nonsignificant increase in tumors at the MTD. The failure of the bioassay to detect this latter type of carcinogen is implicit in Bernstein et al.'s analysis of the hypothetical ideal experiment: the requirement that  $q$  be less than 0.8 specifically excludes them.

Bernstein et al., after noting the experimental absence of

chemicals which produce 100% tumors, then go on to exclude the possibility that such chemicals could appear (if they existed) within the context of their ideal two-group experiment. By arbitrarily setting an upper limit of 98% ( $q \geq 0.02$ ) on the fraction of animals that could possibly get tumors, the upper limit on their estimate of carcinogenic potency then becomes  $3.807/d$ . Taking the ratio of upper ( $3.807/d$ ) and lower ( $1.118/d$ ) limits, they find that the statistically significant values of  $b$  can vary only over a 32-fold range. But by placing an artificial upper limit (98%) on the possible fraction with tumors (and hence, on  $b$ ), they have biased the outcome.

Rieth and Starr examined the arithmetic relationship between  $1/\text{MaxD}$  and the maximum finite value of carcinogenic potency,  $\beta_{\text{max}}$ , for 83 carcinogens chosen from the CPDB (43). How these particular chemicals were selected was not revealed. They found that the mean of the difference  $[\beta_{\text{max}} - (1/\text{MaxD})]$  was  $9.5 \pm 2.2$ , and that this difference is closely tied to the range of doses tested in the bioassay. For vinyl chloride and TCDD, which were tested over 200-fold and 50-fold dose ranges, respectively, the difference between  $\beta_{\text{max}}$  and  $1/\text{MaxD}$  was more than an order of magnitude larger than the mean. The authors concluded that "the doses tested severely and artifactually constrain the estimates of carcinogenic potency that can be derived from the multistage model."

We agree with Rieth and Starr (43) and Bernstein et al. (41) that finite values of  $\beta$  are constrained by both the MaxD and the dose range. But these workers did not take a careful look at whether the relationship between the measured value of  $\beta$  and MaxD is stronger than what would be predicted based on these constraints alone. We address this question elsewhere (50), and report that for chemicals with  $\text{TD}_{50}$  values significant at  $p < 0.01$ , the relationship between  $1/\text{TD}_{50}$  and  $1/\text{MaxD}$  is weaker (has larger variance) for mutagens than for nonmutagens. The fact that there exists a significant difference depending upon mutagenicity, which is an unrelated variable, suggests that the relationship is indeed stronger than what is implied by the constraints alone. Therefore at least a portion of the correlation is nonspurious. Furthermore, we found that the so-called two-dose (zero and MaxD) model of Bernstein et al. does not approximate the actual distribution of  $\beta$  versus  $1/\text{MaxD}$  closely enough to be useful for examining artifacts in the apparent relationship between these two variables. But we did not rule out the possibility, especially for mutagens, that there is little more (or no more) quantitative information to be gained from the relationship between carcinogenic potency and MTD than is already contained in a) the statistical significance level at which the potency is chosen, and b) the fact that chemicals producing 100% tumors at the MTD are rare (50). Figure 5, a plot of  $1/\text{TD}_{50}$  versus  $1/\text{MaxD}$  for chemicals tested by the NCI/NTP and which produced a statistically significant increase in tumors at any site in mice, is reproduced from Goodman et al. (50); the overlap between mutagens and nonmutagens is self-evident.

There is a similar potential bias in the direct comparison of carcinogenicity in humans and rodents as performed by Allen et al. (89). This arises due to the fact that, for most of the chemicals which they analyzed, people were also exposed at high doses. In occupational settings, exposure limits used to be set just below the level at which toxic effects were immediately obvious; to reduce the levels further was considered an unnecessary expense. Chemotherapeutic doses are also close to the maximum tolerated level; lower doses are usually correspondingly less effective.

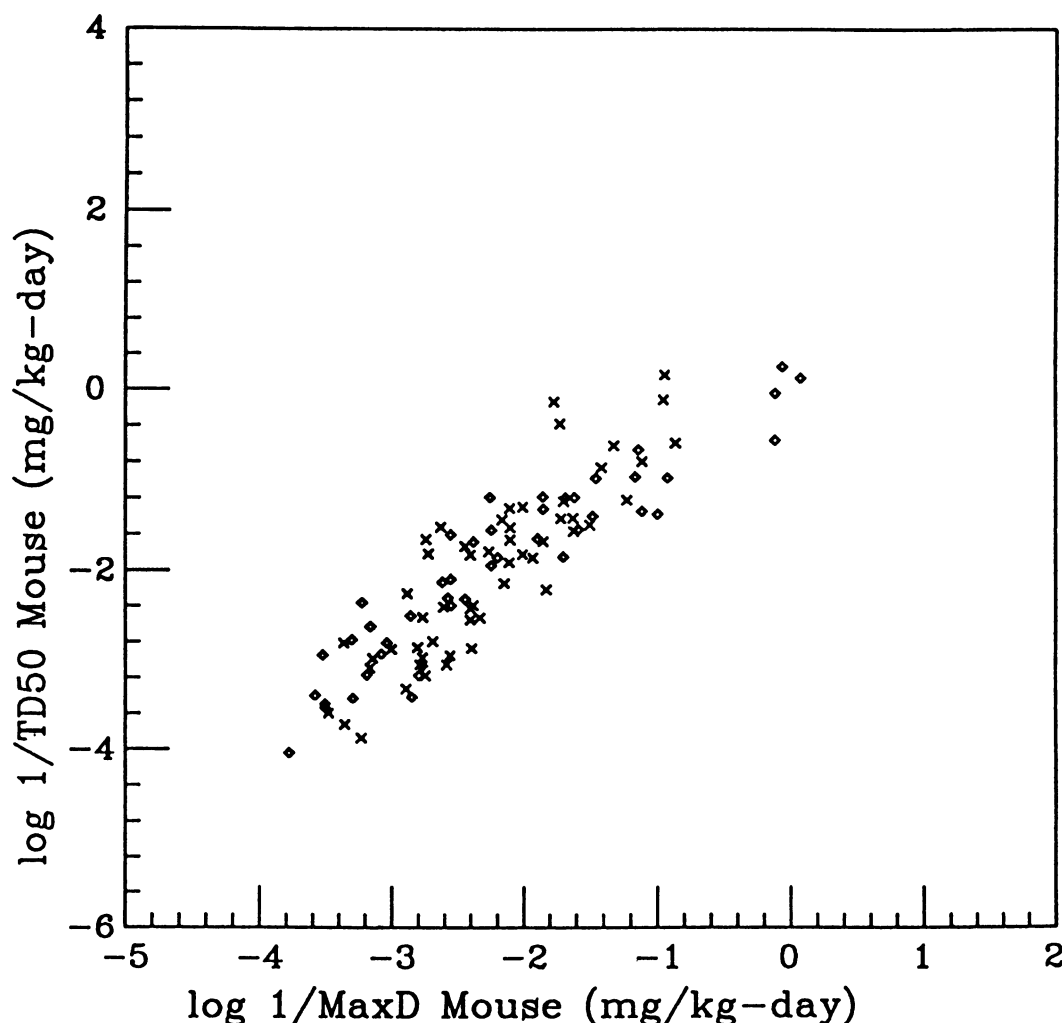


FIGURE 5. Log-log plot of  $1/TD_{50}$  versus  $1/MaxD$  for NCI/NTP chemicals tested in mice. (X) mutagens; ( $\diamond$ ) nonmutagens.  $TD_{50}$  values significant at  $p < 0.01$ . From Goodman et al. (50).

Therefore, any argument that the correlation of carcinogenic potencies for mice and rats is a spurious consequence of the correlation of toxic doses would likewise apply to the correlation of carcinogenic potencies for rodents and humans.

Another bias relates to the choice of chemicals for testing in rodent bioassays. The relative ease of *in vitro* testing for mutagenic and other genotoxic activity makes it unlikely that any new chemical will be considered for mass production if it turns out to be genotoxic. Consequently, fewer new carcinogenicity bioassays are performed on chemicals that are genotoxic in any of the standard *in vitro* tests.

Undoubtedly, there are other biases that have not been identified. However, we believe that it is already clear that a study of the interspecies correlation factor  $K$  cannot be separated from a study of the relationship between toxicity and carcinogenic potency, or of the relationship between carcinogenic potency and activities at the cellular level (including cytotoxicity and genotoxicity). We agree with Clayson (152,153) and others who declare the necessity to improve the carcinogen risk assessment process by increasing and utilizing our understanding of the biological mechanisms of carcinogenesis.

## Genotoxic Versus Nongenotoxic Agents and the Rodent Bioassay

Recently, Ashby and Tennant, expanding upon and corroborating the work of Tennant et al. (154), found that the distribution of tumor sites for agents mutagenic to *Salmonella* is different from those which are not (76). They examined 222 carcinogens which had been tested by the NCI/NTP in both mice and rats. If a chemical was found to be mutagenic in *Salmonella* and also had certain structural attributes which have been associated with mutagenicity (155), it was classified by these authors as genotoxic (+/+). If negative in both, it was classified as nongenotoxic (-/-). The +/+ chemicals produced all tabulated tumors except for seminal vesicle, cholangioma, urinary tract, and lymphatic system. The -/- carcinogens were more restricted in their range, producing tumors at only 15 of the 31 tabulated tumor sites. Benzene, which is a potent genotoxin in most *in vitro* and *in vivo* short-term tests, nevertheless fails to induce point mutations in the *Salmonella* assay and does not have any of the structural features identified by Ashby (155) as predictive of mutagenicity. Thus, benzene is -/- by the above criteria, but Ashby and



Tennant do not include it in the analysis of tumor site distribution because to do so would mask the site differences between truly nongenotoxic carcinogens and genotoxic carcinogens. The liver was the most common target site for both  $+/+$  and  $-/-$  carcinogens, but a  $-/-$  carcinogen was approximately twice as likely as a  $+/+$  carcinogen to cause liver tumors. Of those chemicals which caused tumors only in the mouse liver, 70% were not mutagenic in *Salmonella*. These findings concerning mouse liver tumors were in general agreement with those of Ward et al. (156).

The Science Advisory Board of the U.S. Environmental Protection Agency (EPA) spent a day in August 1987 discussing whether or not induction of mouse liver tumors (and for another reason, rat kidney tumors) provides sufficient evidence that a chemical is a complete carcinogen. One widely held view was that chemicals which produce only mouse liver tumors, particularly in the male B6C3F<sub>1</sub> mouse (which has a high natural incidence of liver tumors and elevated peroxisome levels compared to females and to other strains), may be acting as tumor promoters rather than as complete carcinogens. That hepatic peroxisome proliferators also induce hepatocellular carcinoma is well known (157). The paper by Ashby and Tennant indirectly corroborates the view that most carcinogens which are specific for mouse liver are not primary carcinogens, and that they may act instead to promote the development of neoplasms in pre-initiated cells (76). Gold et al. also found such a distinction: for the 91 mouse liver carcinogens which had been tested for mutagenicity in *Salmonella*, 32% (8/25) of the single-site carcinogens are mutagenic, compared to 56% (37/66) of the multiple-site carcinogens. But a contradictory note is struck by the observation that for the 20 NCI/NTP chemicals which are single-site B6C3F<sub>1</sub> mouse liver carcinogens, tumorigenicity is not strongly correlated with sex: 13 produce tumors in both sexes, 5 in the male only, and 2 in the female only (136). An even weaker correlation with sex was previously found for 26 single-site B6C3F<sub>1</sub> mouse liver carcinogens by a group of NTP scientists (158).

The most important conclusions of Ashby and Tennant (76) were that "screening chemicals for genotoxicity using structural analysis and a minimum number of genotoxicity assays, and use of a reduced cancer bioassay protocol, would enable the detection of trans-species/multiple-site rodent carcinogens. The detection of tissue/sex/species-specific carcinogens can only be achieved by conducting life-time carcinogenicity bioassays according to the present NTP protocol."

## Regulatory Demands and Paradoxes

The fact that about half of the chemicals tested in rodents produce statistically significant tumor increases has important regulatory consequences. It is no longer possible to have a "ban them all" approach. Many synthetic and naturally occurring chemicals which have some measurable carcinogenic activity are important in everyday life. Chloroform is formed in drinking water from the reaction of chlorine (added to curb bacterial growth) with organic contaminants present in most public water supplies. The EPA faces the paradox that it is forced to accept a maximum contaminant level (MCL) for chloroform in public water systems of 100 ppb (159), yet the MCL goal for trichloroethylene is set at zero (160), even though the latter chemical is an order of magnitude less potent a carcinogen than

chloroform. The U.S. Food and Drug Administration (FDA) and the EPA together promulgate inconsistencies in their attempts to limit human exposure to carcinogens; the FDA accepts peanut butter with 20 ppb total aflatoxins (including B<sub>1</sub>) and have proposed that the tolerance be lowered to 15 ppb (161), even though aflatoxin B<sub>1</sub> is at least  $10^5$  times more potent than trichloroethylene.

The EPA justifies its use of the surface area correction factor (which the FDA does not use) on the ground that it is more "conservative." We do not contest this desire to be conservative, but suggest that the factors be explicitly included. We have argued that  $K$  should be unity within an uncertainty factor of 20 either way (52). If it is desired to be conservative, one may take 20, so long as it is clearly recognized that there are many chemicals for which this is high. It should also be recognized that for many chemicals  $K$  would be 1, for some  $K$  would be 1/20, and for any untested chemical  $K$  may approach zero.

The interspecies potency factors discussed here are intended to be used in numerical risk estimates. It must be recognized that many distinguished scientists do not accept the use by ourselves and by government agencies of numerical values for interspecies factors in calculating a risk. For example, Ames et al. use the term "possible hazard" for estimates of human carcinogenicity based on rodent bioassays and do not write the word "risk" unless data is available from human exposures (4). They produced a human exposure dose/rodent potency dose (HERP) index, in which the human lifetime daily dose is expressed as a percentage of the rodent TD<sub>50</sub>, where both are in units of mg/kg-day. For  $K = 1$ , the HERP divided by  $100/\ln(2)$  becomes what we would call the risk, and the order of chemicals listed by HERP value is the same as if they were listed by risk. In the text, Ames et al. use their list to suggest that exposures associated with a small HERP value can be ignored compared to substances with a larger one. This use can only be valid if it is assumed that the HERP index is related to risk. Thus all the assumptions, restrictions, and qualifications appropriate to the interspecies comparison of carcinogenic potencies also apply whenever the HERP is used.

Doll and Peto emphasized the uncertainty in quantitative estimation of human risk from animal data and suggested that "priority setting" should replace "risk assessment" (84). Specifically, they recommended that a chemical's potency in each test (including long-term carcinogenicity bioassays and short-term *in vitro* tests) be multiplied by an estimate of human exposure to yield an index of human hazard according to that one test. Chemicals that appeared high on any index would be considered prime candidates for regulatory action. One appealing aspect of this recommendation is that it insures that no chemical which is potent by any criterion is overlooked. We note that the above-mentioned index of Ames et al. is an example of the procedure suggested by Doll and Peto when the test under consideration is the rodent bioassay.

## Suggestions for Future Work

An important issue concerning high-dose/low-dose extrapolation might be resolved by careful documentation of the dose-response for cell proliferation. If it is true that cell proliferation is responsible for carcinogenicity at toxic doses, as is now widely believed, then a corollary hypothesis is immediately apparent: that chemicals which are not genotoxic in a given target organ and which do not cause local cell proliferation will not cause cancer

in that organ. This is a powerful notion; one with unprecedented power to change the face of carcinogen risk assessment.

We wish to emphasize the need for adequate rodent bioassays, in which the route of administration is matched to the human exposure route, for chemicals already known to be carcinogenic to humans. In the absence of good animal data for all known human carcinogens, refinement of the methodology for predicting potency in humans from potency in animals is seriously hindered. Likewise, there is a dearth of knowledge as to the potencies of known human carcinogens. For some chemicals, the necessary quantitative information could be derived from completed or ongoing epidemiological studies. If epidemiologists could be persuaded of the importance of obtaining potency estimates, then more studies would be designed so as to reveal the effects at different exposure levels. In those epidemiological studies in which no statistically significant increase in neoplasms are found, we encourage the determination of upper limits to carcinogenic potency.

We thank K. S. Crump, D. W. Gaylor, L. S. Gold, and many others who have sent us their latest views, data, and preprints. We are especially grateful to J. K. Haseman for commenting extensively on an earlier version of the manuscript. R. Wilson was supported by an intergovernmental exchange agreement with the FDA. G. Goodman was supported by gifts to Harvard University from Bristol-Myers, Gillette, and Monsanto Corporations, and by cooperative agreement CR812699 from the EPA to the Programs in Environmental Health and Public Policy at the Harvard School of Public Health.

## REFERENCES

- Innes, J. R. M., Ulland, B. M., Valerio, M. G., Petrucelli, L., Fishbein, L., Hart, E. R., Pallota, A. J., Bates, R. R., Falk, H. L., Gart, J. J., Klein, M., Mitchell, I., and Peters, J. Bioassay of pesticides and industrial chemicals for tumorigenicity in mice: a preliminary note. *J. Natl. Cancer Inst.* 42: 1101-1114 (1969).
- Epstein, S. S., and Swartz, J. B. Carcinogenic risk assessment. *Science* 240: 1043-1045 (1988).
- Ames, B. N., Magaw, R., and Gold, L. S. Ranking possible carcinogenic hazards. *Science* 236: 271-280 (1987).
- Ames, B. N., and Gold, L. S. Response to 'Carcinogenic risk assessment'. *Science* 240: 1045-1047 (1988).
- Office of Science and Technology Policy. Chemical carcinogens: a review of the science and its associated principles. *Fed. Reg.* 50: 10371-10442 (1985).
- Haseman, J. K., and Huff, J. E. Species correlation in long-term carcinogenicity studies. *Cancer Lett.* 37: 125-132 (1987).
- Haseman, J. K., Huff, J. E., Zeiger, E., and McConnell, E. E. Comparative results of 327 chemical carcinogenicity studies. *Environ. Health Perspect.* 74: 229-235 (1987).
- Byrd, D. M., Crouch, E. A. C., and Wilson, R. Do mouse liver tumors predict rat tumors? A study of concordance between tumors induced at different sites in rats and mice. In: *Mouse Liver Carcinogenesis: Mechanisms and Species Comparisons* (D. E. Stevenson, Ed.), Alan R. Liss, New York, 1990, pp. 19-41.
- Fears, T. R., Tarone, R. E., and Chu, K. C. False-positive and false negative rates for carcinogenicity screens. *Cancer Res.* 37: 1941-1945 (1977).
- Gart, J. J., Chu, K. C., and Tarone, R. E. Statistical issues in interpretation of chronic bioassays for carcinogenicity. *J. Natl. Cancer Inst.* 62: 957-974 (1979).
- Haseman, J. K. A reexamination of false positive rates for carcinogenicity studies. *Fundam. Appl. Toxicol.* 12: 793-804 (1983).
- Hutt, P. B. Use of quantitative risk assessment in regulatory decision making under Federal health and safety statutes. In: *Risk Quantitation and Regulatory Policy*. Banbury Report 19 (D. G. Hoel, R. A. Merrill, and F. P. Perera, Eds.), Cold Spring Harbor Laboratory, New York, 1985, pp. 15-29.
- Scientific Committee of the Food Safety Council. Quantitative risk assessment. *Fd. Cosmet. Toxicol.* 18: 711-734 (1980).
- Gregory, A. R. Species comparisons in evaluating carcinogenicity from animal studies. *Regul. Toxicol. Pharmacol.* 8: 160-190 (1988).
- Dybing, E. Predictability of human carcinogenicity from animal studies. *Regul. Toxicol. Pharmacol.* 6: 399-415 (1986).
- Ramel, C. Short-term testing—are we looking at wrong endpoints? *Mutat. Res.* 205: 13-24 (1988).
- Weinstein, I. B. The origins of human cancer: molecular mechanisms of carcinogenesis and their implications for cancer prevention and treatment—twenty-seventh G.H.A. Clowes Memorial Lecture. *Cancer Res.* 48: 4135-4143 (1988).
- Weinstein, I. B. The relevance of tumor promotion and multistage carcinogenesis to risk assessment. In: *Risk Quantitation and Regulatory Policy*. Banbury Report 19 (D. G. Hoel, R. A. Merrill, and F. P. Perera, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1985, pp. 231-238.
- International Commission for Protection against Environmental Mutagens and Carcinogens (ICPEMC). Publication No. 16, Testing for mutagens and carcinogens; the role of short-term genotoxicity assays. *Mutat. Res.* 205: 3-12 (1988).
- Shelby, M. D. The genetic toxicity of human carcinogens and its implications. *Mutat. Res.* 204: 3-15 (1988).
- McCann, J., Choi, E., Yamasaki, E., and Ames, B. N. Detection of carcinogens as mutagens in the *Salmonella*/microsome test: assay of 300 chemicals. *Proc. Natl. Acad. Sci. U.S.A.* 72: 5135-5139 (1975).
- Purchase, I. F. H., Longstaff, E., Ashby, J., Styles, J. A., Anderson, D., Lefevre, P. A., and Westwood, F. R. Evaluation of six short term tests for detecting organic chemical carcinogens and recommendations for their use. *Nature* 264: 624-627 (1976).
- Tennant, R. W., Margolin, B. H., Shelby, M. D., Zeiger, E., Haseman, J. K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B., and Minor, R. Prediction of chemical carcinogenicity in rodents from *in vitro* genetic toxicity assays. *Science* 236: 933-941 (1987).
- Zeiger, E. Carcinogenicity of mutagens: predictive capability of the *Salmonella* mutagenesis assay for rodent carcinogenicity. *Cancer Res.* 47: 1287-1296 (1987).
- Nesnow, S., and Bergman, H. An analysis of the Gene-Tox Carcinogen Data Base. *Mutat. Res.* 205: 237-253 (1988).
- Meselson, M., and Russell, K. Comparisons of carcinogenic and mutagenic potency. In: *Origins of Human Cancer* (H. H. Hiatt, J. D. Watson, and J. A. Weinstein, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 1473-1481.
- Jones, C. A., Marlino, P. J., Lijinsky, W., and Huberman, E. The relationship between the carcinogenicity and mutagenicity of nitrosamines in a hepatocyte-mediated mutagenicity assay. *Carcinogenesis* 2: 1075-1077 (1981).
- Bartsch, H., Tomatis, L., and Malaveille, C. Qualitative and quantitative comparisons between mutagenic and carcinogenic activities of chemicals. In: *Mutagenicity: New Horizons in Genetic Toxicology* (J. A. Heddle, Ed.), Academic Press, New York, 1982, pp. 35-72.
- McCann, J., Gold, L. S., Horn, L., McGill, R., Graedel, T. E., and Kaldor, J. Statistical analysis of *Salmonella* test data and comparison to results of animal cancer tests. *Mutat. Res.* 205: 183-195 (1988).
- Parodi, S., Taningher, M., and Santi, L. Utilization of the quantitative component of positive and negative results of short-term tests. *Mutat. Res.* 205: 283-294 (1988).
- Shelby, M. D., and Stasiewicz, S. Chemicals showing no evidence of carcinogenicity in long-term, two-species rodent studies: the need for short-term test data. *Environ. Mutagen.* 6: 871-878 (1984).
- Waters, M. D., Bergman, H. B., and Nesnow, S. The genetic toxicology of Gene-Tox non-carcinogens. *Mutat. Res.* 205: 139-182 (1988).
- Calabrese, E. J. Comparative biology of test species. *Environ. Health Perspect.* 77: 55-62 (1988).
- Freireich, E. J., Gehan, E. A., Rall, D. P., Schmidt, L. H., and Skipper, H. E. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. *Cancer Chemother. Rep.* 50: 219-244 (1966).
- Rall, D. P. Difficulties in extrapolating the results of toxicity studies in laboratory animals to man. *Environ. Res.* 2: 360-367 (1969).
- Krasovskij, G. N. Species and sex differences in sensitivity to toxic substances. In: *Methods Used in the USSR for Establishing Biologically Safe Levels of Toxic Substances*. World Health Organization, Geneva, 1975, pp. 109-125.
- Crouch, E. A. C. Uncertainties in interspecies extrapolations of carcinogenicity. *Environ. Health Perspect.* 50: 321-327 (1983).

38. Travis, C. C., White, R. K., and Ward, R. C. Interspecies extrapolation of pharmacokinetics. *J. Theor. Biol.* 142(3): 285-304 (1990).
39. Crouch, E., and Wilson, R. Interspecies comparison of carcinogenic potency. *J. Toxicol. Environ. Health* 5: 1095-1118 (1979).
40. Peto, R., Pike, M. C., Bernstein, L., Gold, L. S., and Ames, B. N. The TD<sub>50</sub>: a proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. *Environ. Health Perspect.* 58: 1-8 (1984).
41. Bernstein, L., Gold, L. S., Ames, B. N., Pike, M. C., and Hoel, D. G. Some tautologous aspects of the comparison of carcinogenic potency in rats and mice. *Fundam. Appl. Toxicol.* 5: 79-86 (1985).
42. Crouch, E., Wilson, R., and Zeise, L. Tautology or not tautology? *J. Toxicol. Environ. Health* 20: 1-10 (1987).
43. Rieth, J. P., and Starr, T. B. Experimental design constraints on carcinogenic potency estimates. *J. Toxicol. Environ. Health* 27: 287-296 (1989).
44. Staffa, J. A., and Mehlman, M. A. Innovations in Cancer Risk Assessment (ED01 Study). Pathotox Publishers, Park Forest South, IL, 1979.
45. Carlborg, F. W. 2-Acetylaminofluorene and the Weibull model. *Fd. Cosmet. Toxicol.* 19: 367-371 (1981).
46. Peto, R., Gray, R., Brantom, P., and Grasso, P. Nitrosamine carcinogenesis in 5120 rodents: chronic administration of sixteen different concentrations of NDEA, NDMA, NPYR and NPIP in the water of 4400 inbred rats with parallel studies on NDEA alone of the effect of age of starting (3, 6, or 20 weeks) and of species (rats, mice or hamsters). In: *N-Nitroso Compounds: Occurrence, Biological Effects and Relevance to Human Cancer* (I. K. O'Neill, R. C. von Borstel, C. T. Miller, J. Long, and H. Bartsch, Eds.), IARC Scientific Publications No. 57, International Agency for Research on Cancer, Lyon, 1984, pp. 627-665.
47. Ito, N., Shirai, T., Fukushima, S., and Hirose, M. Dose-response study of urinary bladder carcinogenesis in rats by N-butyl-N-(4-hydroxybutyl)nitrosamine. *J. Cancer Res. Clin. Oncol.* 108: 169-173 (1984).
48. Swenberg, J. A. Chemically-induced cell proliferation as a criterion in selecting doses for long-term bioassays. In: *Chemically Induced Cell Proliferation: Implications for Risk Assessment* (B. E. Butterworth, T. J. Slaga, W. Farland, and M. McClain, Eds.), Wiley-Liss, New York, 1991, pp. 245-251.
49. Cohen, S. M. Cell proliferation and bladder tumor promotion. In: *Chemically Induced Cell Proliferation: Implications for Risk Assessment* (B. E. Butterworth, T. J. Slaga, W. Farland, and M. McClain, Eds.), Wiley-Liss, New York, 1991, pp. 347-355.
50. Goodman, G., Shlyakhter, A., and Wilson, T. The relationship between carcinogenic potency and maximum tolerated dose is similar for mutagens and non-mutagens. In: *Chemically Induced Cell Proliferation: Implications for Risk Assessment* (B. E. Butterworth, T. J. Slaga, W. Farland, and M. McClain, Eds.), Wiley-Liss, New York, 1991, pp. 501-516.
51. Zeise, L., Wilson, R., and Crouch, E. A. C. Dose-response relationships for carcinogens: a review. *Environ. Health Perspect.* 73: 259-308 (1987).
52. Crouch, E., and Wilson, R. Regulation of carcinogens. *Risk Anal.* 1: 47-57 (1981).
53. Crouch, E. A. C. Carcinogenic risk assessment—the consequences of believing models. In: *Organ and Species Specificity in Chemical Carcinogenesis* (R. Langenbach, S. Newnow, and J. M. Rice, Eds.), Plenum Press, New York, 1983, pp. 653-665.
54. Peto, R. Detection of risk of cancer to man. *Proc. R. Soc. Lond. B.* 205: 111-120 (1979).
55. Armitage, P., and Doll, R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer* 8: 1-12 (1954).
56. Armitage, P., and Doll, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br. J. Cancer* 11: 161-169 (1957).
57. Cook, P. J., Doll, R., and Fellingham, S. A. A mathematical model for the age distribution of cancer in man. *Int. J. Cancer* 4: 93-112 (1969).
58. Cohen, B. L. The cancer risk from low-level radiation. *Health Phys.* 39: 659-678 (1980).
59. Rall, D. P. Species differences in carcinogenesis testing. In: *Origins of Human Cancer* (H. H. Hiatt, J. D. Watson, and J. A. Winstein, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 1383-1390.
60. Travis, C. C. Interspecies extrapolations in risk analysis. *Toxicology* 47: 3-13 (1987).
61. Anderson, M. W., Hoel, D. G., and Kaplan, N. L. A general scheme for the incorporation of pharmacokinetics in low-dose risk estimation for chemical carcinogenesis: example—vinyl chloride. *Toxicol. Appl. Pharmacol.* 55: 154-161 (1980).
62. Yamasaki, E., and Ames, B. N. Concentration of mutagens from urine by adsorption with the nonpolar resin XAD-2: cigarette smokers have mutagenic urine. *Proc. Natl. Acad. Sci. U.S.A.* 74: 3555-3559 (1977).
63. Hoel, D. G., Haseman, J. K., Hogan, M. D., Huff, J., and McConnell, E. E. The impact of toxicity on carcinogenicity studies: implications for risk assessment. *Carcinogenesis* 9: 2045-2052 (1988).
64. Moolgavkar, S. H., and Knudson, A. G., Jr. Mutation and cancer: a model for human carcinogenesis. *J. Natl. Cancer Inst.* 66: 1037-1052 (1981).
65. Moolgavkar, S. H. Carcinogenesis modeling: from molecular biology to epidemiology. *Ann. Rev. Public Health* 7: 151-169 (1986).
66. Greenfield, R. E., Ellwein, L. B., and Cohen, S. M. A general probabilistic model of carcinogenesis: analysis of experimental urinary bladder cancer. *Carcinogenesis* 5: 437-445 (1984).
67. Arnold, D. L., Moodie, C. A., Grice, H. C., Charbonneau, S., Stavric, B., Collins, B. F., McGuire, P. F., Zawadzka, Z. Z., and Munro, I. C. Long term toxicity of *ortho*-toluenesulfonamide and sodium saccharin in the rat. *Toxicol. Appl. Pharmacol.* 52: 113-152 (1980).
68. Ellwein, L. B., and Cohen, S. M. A cellular dynamics model of experimental bladder cancer: analysis of the effect of sodium saccharin in the rat. *Risk Anal.* 8: 215-221 (1988).
69. Weisburger, J. H., and Williams, G. M. The distinct health risk analyses required for genotoxic carcinogens and promoting agents. *Environ. Health Perspect.* 50: 233-245 (1983).
70. Potter, V. R. A new protocol and its rationale for the study of initiation and promotion of carcinogenesis in rat liver. *Carcinogenesis* 2: 1375-1379 (1981).
71. Potter, V. R. Alternative hypotheses for the role of promotion in chemical carcinogenesis. *Environ. Health Perspect.* 50: 139-148 (1983).
72. Kunz, H. W., Tennekens, H. A., Port, R. E., Schwartz, M., Lorke, D., and Schade, G. Quantitative aspects of chemical carcinogenesis and tumor promotion in liver. *Environ. Health Perspect.* 50: 113-122 (1983).
73. Barbacid, M. *ras* genes. *Annu. Rev. Biochem.* 56: 779-827 (1987).
74. Farr, C. J., Saiki, R. K., Erlich, H. A., McCormick, F., and Marshall, C. J. Analysis of *ras* gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. *Proc. Natl. Acad. Sci. U.S.A.* 85: 1629-1633 (1988).
75. Ashby, J. Screening chemicals for mutagenicity: practices and pitfalls. In: *Mutagenicity: New Horizons in Genetic Toxicology* (J. A. Heddle, Ed.), Academic Press, New York, 1982, pp. 1-33.
76. Ashby, J., and Tennant, R. W. Chemical structure, *Salmonella* mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutat. Res.* 204: 17-115 (1988).
77. Farber, E. The multistep nature of cancer development. *Cancer Res.* 44: 5463-5474 (1984).
78. Fischer, S. M., Floyd, R. A., and Copeland, E. S. Oxy radicals in carcinogenesis—a chemical pathology study section workshop. *Cancer Res.* 48: 3882-3887 (1988).
79. Higgingson, J. Cancer and environment. *Science* 205: 1363-1366 (1979).
80. Haenszel, W., and Kurihara, M. Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J. Natl. Cancer Inst.* 40: 43-68 (1968).
81. Cairns, J. The origin of human cancers. *Nature* 289: 353-357 (1981).
82. German, J., Bloom, D., and Passarge, E. Bloom's syndrome XI. Progress report for 1983. *Clin. Genet.* 25: 166-174 (1984).
83. Cairns, J. In: *The Causes of Cancer*. Oxford University Press, New York, 1981, p. 1216.
84. Doll, R., and Peto, R. *The Causes of Cancer*. Oxford University Press, New York, 1981.
85. Knudson, A. G., Jr. Hereditary cancer, oncogenes, and antioncogenes. *Cancer Res.* 45: 1437-1443 (1985).
86. Higgingson, J. Population studies in cancer. *Acta. Un. Cancer* 16: 1667-1670 (1960).
87. Wynder, E. L., and Gori, G. B. Contribution of the environment to cancer incidence: an epidemiologic exercise. *J. Natl. Cancer Inst.* 58: 825-832 (1977).
88. Calabrese, E. J. *Principles of Animal Extrapolation*. John Wiley and Sons, New York, 1983, pp. 499-527.
89. Allen, B. C., Crump, K. S., and Shipp, A. M. Correlation between carcinogenic potency of chemicals in animals and humans. *Risk Anal.* 8: 531-544 (1988).
90. Silverberg, S. G. The autopsy and cancer. *Arch. Pathol. Lab. Med.* 108: 476-478 (1984).

91. Roe, F. J. C. Opinions on animal selection for the assessment of carcinogenicity. In: *Human Risk Assessment—The Role of Animal Selection and Extrapolation* (M. V. Rolloff, Ed.), Taylor and Francis, Philadelphia, 1987, pp. 31–44.
92. Roe, F. J. C. Testing for carcinogenicity and the problem of pseudocarcinogenicity. *Nature* 303: 657–658 (1983).
93. Conybeare, G. Effect of quality and quantity of diet on survival and tumour incidence in outbred Swiss mice. *Fd. Cosmet. Toxicol.* 18: 65–72 (1980).
94. Newberne, P. M., and Conner, M. W. The role of diet and nutrition in safety evaluation of drugs and other chemicals. In: *Human Risk Assessment—The Role of Animal Selection and Extrapolation* (M. V. Rolloff, Ed.), Taylor and Francis, Philadelphia, 1987, pp. 65–79.
95. Byrd, D. M., Crouch, E. A. C., and Wilson, R. Organ specificity of rodent tumor induction. Presented at the Society for Risk Analysis Annual Meeting, November, 1988, Washington, DC.
96. Busby, W. F., and Wogan, G. N. Aflatoxins. In: *Chemical Carcinogens*, Vol. 2 (C. S. Searle, Ed.), ACS Monograph 182, American Chemical Society, Washington, DC, 1984, pp. 945–1136.
97. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Supplement 7: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42. International Agency for Research on Cancer, Lyon, 1987.
98. Yeh, F.-S., Yu, M. C., Mo, C.-C., Luo, S., Tong, M. S., and Henderson, B. E. Hepatitis B virus, aflatoxins, and hepatocellular carcinoma in southern Guangxi, China. *Cancer Res.* 49: 2506–2509 (1989).
99. Maltoni, C., and Lefemine, G. Carcinogenicity bioassays of vinyl chloride. I. Research plan and early results. *Environ. Res.* 7: 387–405 (1974).
100. Maltoni, C., and Lefemine, G. Carcinogenicity bioassays of vinyl chloride: current results. *Ann. N.Y. Acad. Sci.* 246: 195–218 (1975).
101. Creech, J. L., and Johnson, M. N. Angiosarcoma of liver in the manufacture of polyvinyl chloride. *J. Occup. Med.* 16: 150–151 (1974).
102. Hammond, E. C. and Horn, D. Smoking and death rates - report on forty-four months of follow-up of 187,783 men. *Ca-A Cancer J. Clin.* 38: 28–58 (1988).
103. Nicholson, W. J., Henneberger, P. K., and Seidman, H. Occupational hazards in the VC-PC industry. *Prog. Clin. Biol. Res.* 141: 155–175 (1984).
104. Tabershaw, I. R., and Gaffey, W. R. Mortality study of workers in the manufacture of vinyl chloride and its polymers. *J. Occup. Med.* 16: 509–518 (1974).
105. Pirastu, R., Comba, P., Reggiani, A., Foa, V., Masina, A., and Maltoni, C. Mortality from liver disease among Italian vinyl chloride monomer/polyvinyl chloride manufacturers. *Am. J. Ind. Med.* 17: 155–161 (1990).
106. Tomatis, L., Partensky, C., and Montesano, R. The predictive value of mouse liver tumour induction in carcinogenicity testing—a literature survey. *Int. J. Cancer* 12: 1–20 (1973).
107. Tomatis, L., Agthe, C., Bartsch, H., Huff, J., Montesano, R., Saracci, R., Walker, E., and Wilbourn, J. Evaluation of the carcinogenicity of chemicals: a review of the Monograph Program of the International Agency for Research on Cancer (1971 to 1977). *Cancer Res.* 38: 877–885 (1978).
108. Maltoni, C., Ciliberti, A., Cotti, G., Conti, B., and Belpoggi, F. Benzene, an experimental multipotential carcinogen: results of the long-term bioassays performed at the Bologna Institute of Oncology. *Environ. Health Perspect.* 82: 109–124 (1989).
109. Huff, J. E., Haseman, J. K., DeMarini, D. M., Eustis, S., Maronpot, R. R., Peters, A. C., Persing, R. L., Chrisp, C. E., and Jacobs, A. C. Multiple-site carcinogenicity of benzene in Fischer 344 rats and B6C3F<sub>1</sub> mice. *Environ. Health Perspect.* 82: 125–163 (1989).
110. Rinsky, R. A., Smith, A. B., Hornung, R., Filloon, T. G., Young, R. J., Okun, A. H., and Landrigan, P. J. Benzene and leukemia: an epidemiologic risk assessment. *N. Engl. J. Med.* 316: 1044–1050 (1987).
111. Anderson, E. L., and the Carcinogen Assessment Group of the U.S. Environmental Protection Agency. Quantitative approaches in use to assess cancer risk. *Risk Anal.* 3: 277–295 (1983).
112. Bannasch, P. Sequential cellular changes during chemical carcinogenesis. *J. Cancer Res. Clin. Oncol.* 108: 11–22 (1984).
113. Bannasch, P., Griesemer, R. A., Anders, F., Becker, R., Cabral, J. R., Della Porta, G., Feron, V. J., Henschler, D., Ito, N., Kroes, R., Magee, P. N., Montesano, R., Napalkov, N. P., Nesnow, S., Rao, G. N., and Wilbourn, J. Early preneoplastic lesions. In: *Long-Term and Short-Term Assays for Carcinogens: A Critical Appraisal*, IARC Scientific Publications No. 83, International Agency for Research on Cancer, Lyon, 1986, pp. 85–101.
114. Huff, J. E., Eustis, S. L., and Haseman, J. K. Occurrence and relevance of chemically induced benign neoplasms in long-term carcinogenicity studies. *Cancer Metastas. Rev.* 8: 1–21 (1989).
115. Wilbourn, J., Haroun, L., Heseltine, E., Kaldor, J., Partensky, C., and Vainio, H. Response of experimental animals to human carcinogens: an analysis based upon the IARC Monographs programme. *Carcinogenesis* 7: 1853–1863 (1986).
116. Ishinishi, N., Yamamoto, A., Hisanaga, A., and Inamasu, T. Tumorigenicity of arsenic trioxide to the lung in Syrian golden hamsters by intermittent instillations. *Cancer Lett.* 21: 141–147 (1983).
117. Pershagen, G., Nordberg, G., and Björklund, N. -E. Carcinomas of the respiratory tract in hamsters given arsenic trioxide and/or benzo[a]pyrene by the pulmonary route. *Environ. Res.* 34: 227–241 (1984).
118. U.S. Environmental Protection Agency. Assessment of Health Effects of Benzene Germane to Low-Level Exposure. EPA-600/1-78-061, EPA, Washington, DC, 1978.
119. Snyder, C. A., Goldstein, B. D., Sellakumar, A. R., Bromberg, I., Laski, S., and Albert, R. E. The inhalation toxicology of benzene: incidence of hematopoietic neoplasms and hematotoxicity in AKR/J and C57BL/6J mice. *Toxicol. Appl. Pharmacol.* 54: 323–331 (1980).
120. National Toxicology Program. Toxicology and Carcinogenesis Studies of Benzene in F344/N Rats and B6C3F<sub>1</sub> Mice (Gavage Studies). Technical Report 289, U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD, 1986.
121. Freedman, D. A., and Zeisel, H. From mouse to man: the quantitative assessment of cancer risks. *Stat. Sci.* 3: 3–56 (1988).
122. MacMahon, B., Monson, R. R., Wang, H. H., and Zheng, T. A second follow-up of mortality in a cohort of pesticide applicators. *J. Occup. Med.* 30: 429–432 (1988).
123. Enterline, P. E., and Viren, J. Epidemiologic evidence for an association between gasoline and kidney cancer. *Environ. Health Perspect.* 62: 303–312 (1985).
124. Wen, C. P. Tsai, S. P., Moffit, K. B., Bondy, M., and Gibson, R. L. Epidemiological studies of the role of gasoline (hydrocarbon) exposure in kidney cancer risk. In: *Renal Effects of Petroleum Hydrocarbons* (M. A. Mehlman, G. P. Hemstreet, J. J. Thorpe, and N. K. Weaver, Eds.), Princeton Scientific, Princeton, NJ, 1984, pp. 245–257.
125. Thomas, T. L., Decoufle, P., and Moure-Eraso, R. Mortality among workers employed in petroleum refining and petrochemical plants. *J. Occup. Med.* 22: 97–103 (1980).
126. MacFarland, H. N., Ulrich, C. E., Holdsworth, C. E., Kitchen, D. N., Halliwell, W. H., and Blum, S. C. A chronic inhalation study with unleaded gasoline vapor. *J. Am. Coll. Toxicol.* 3: 231–248 (1984).
127. Ennever, F. K., Noonan, T. J., and Rosenkranz, H. S. The predictivity of animal bioassays and short-term genotoxicity tests for carcinogenicity and non-carcinogenicity to humans. *Mutagenesis* 2: 73–78 (1987).
128. Goodman, G., and Wilson, R. W. Quantitative prediction of human cancer risk from rodent carcinogenic potencies: a closer look at the epidemiological evidence for some chemicals not definitively carcinogenic in humans. *Regul. Toxicol. Pharmacol.*, in press.
129. Gold, L. S., Sawyer, C. B., Magaw, R., Backman, G. M., de Veciana, M., Levinson, R., Hooper, N. K., Havender, W. R., Bernstein, L., Peto, R., Pike, M. C., and Ames, B. N. A carcinogenic potency database of the standardized results of animal bioassays. *Environ. Health Perspect.* 58: 9–319 (1984).
130. Gold, L. S., de Veciana, M., Backman, G. M., Magaw, R., Lopipero, P., Smith, M., Blumenthal, M., Levinson, R., Bernstein, L., and Ames, B. N. Chronological supplement to the carcinogenic potency database: standardized results of animal bioassays published through December 1982. *Environ. Health Perspect.* 67: 161–200 (1986).
131. Gold, L. S., Slone, T. H., Backman, G. M., Magaw, R., Da Costa, M., Lopipero, P., Blumenthal, M., and Ames, B. N. Second chronological supplement to the carcinogenic potency database: standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. *Environ. Health Perspect.* 74: 237–329 (1987).
132. Hottendorf, G. H. Species differences in toxic lesions. In: *Human Risk Assessment—The Role of Animal Selection and Extrapolation* (M. V. Rolloff, Ed.), Taylor and Francis, Philadelphia, 1987, pp. 87–95.
133. Sontag, J. M. Aspects in carcinogen bioassay. In: *Origins of Human Cancer* (H. H. Hiatt, J. D. Watson, and J. A. Winstein, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 1327–1338.
134. Sontag, J. M., Page, N., and Saffioti, U. Guidelines for Carcinogen Bioassay in Small Rodents. DHEW Publication No. (NIH) 76-801, U.S. Dept. of Health, Education and Welfare, Washington, DC, 1976.
135. Purchase, I. F. H. Inter-species comparisons of carcinogenicity. *Br. J. Cancer* 41: 454–468 (1980).

136. Gold, L. S., Bernstein, L., Magaw, R., and Slone, T. H. Interspecies extrapolation in carcinogenesis: prediction between rats and mice. *Environ. Health Perspect.* 81: 211-219 (1989).
137. Young, S. S. A blind reanalysis of a random subset of NCI bioassay studies: agreement between rats and mice. Introduction to Symposium. *Fundam. Appl. Toxicol.* 12: 189-190 (1989).
138. Young, S. S. A blind reanalysis of a random subset of NCI bioassay studies: agreement between rats and mice. Discussion. *Fundam. Appl. Toxicol.* 12: 232-241 (1989).
139. Sanathanan, L. P., Lin, C. T., and Carr, R. N. A blind reanalysis of a random subset of NCI bioassay studies: do rats predict mice? *Fundam. Appl. Toxicol.* 12: 191-201 (1989).
140. Louis, T. A. Analysis of a random sample of 2-year carcinogen bioassays from the NCI data base. *Fundam. Appl. Toxicol.* 12: 222-231 (1989).
141. Bickis, M., and Krewski, D. Statistical issues in the analysis of the long-term carcinogenicity bioassay in small rodents: an empirical evaluation of statistical decision rules. *Fundam. Appl. Toxicol.* 12: 202-221 (1989).
142. Haseman, J. K. Use of statistical decision rules for evaluating laboratory animal carcinogenicity studies. *Fundam. Appl. Toxicol.* 14: 637-648 (1990).
143. Metzger, B., Crouch, E., and Wilson, R. On the relationship between carcinogenicity and acute toxicity. *Risk Anal.* 9: 169-177 (1989).
144. Rieth, J. P., and Starr, T. B. Chronic bioassays: relevance to quantitative risk assessment of carcinogens. *Regul. Toxicol. Pharmacol.* 10: 160-173 (1989).
145. Gaylor, D. W., and Chen, J. J. Relative potency of chemical carcinogens in rodents. *Risk Anal.* 6: 283-290 (1986).
146. Chen, J. J., and Gaylor, D. W. Carcinogenic risk assessment: comparison of estimated safe doses for rats and mice. *Environ. Health Perspect.* 72: 305-309 (1987).
147. Clayson, D. B., and Krewski, D. International Commission for Protection against Environmental Mutagens and Carcinogens (ICPEMC) Publication No. 3, The concept of negativity in experimental carcinogenesis. *Mutat. Res.* 167: 233-240 (1986).
148. Zeise, L., Wilson, R., and Crouch, E. Use of acute toxicity to estimate carcinogenic risk. *Risk Anal.* 4: 187-199 (1984).
149. Zeise, L., Crouch, E. A. C., and Wilson, R. Reply to comments: on the relationship of toxicity and carcinogenicity. *Risk Anal.* 5: 265-270 (1985).
150. Zeise, L., Crouch, E. A. C., and Wilson, R. A possible relationship between toxicity and carcinogenicity. *J. Am. Coll. Toxicol.* 5: 137-151 (1986).
151. Parodi, S., Taningher, M., Boero, P., and Santi, L. Quantitative correlations amongst alkaline DNA fragmentation, DNA covalent binding, mutagenicity in the Ames test and carcinogenicity. *Mutat. Res.* 93: 1-24 (1982).
152. Clayson, D. B. International Commission for Protection against Environmental Mutagens and Carcinogens (ICPEMC) Publication No. 13, The need for biological risk assessment in reaching decisions about carcinogens. *Mutat. Res.* 185: 243-269 (1987).
153. Clayson, D. B. Needs for biological risk assessment in interspecies extrapolation. *Environ. Health Perspect.* 77: 93-97 (1988).
154. Tennant, R. W., Stasiewicz, S., and Spalding, J. W. Comparison of multiple parameters of rodent carcinogenicity and *in vitro* genetic toxicity. *Environ. Mutagen.* 8: 205-227 (1986).
155. Ashby, J. Fundamental structural alerts to potential carcinogenicity or non-carcinogenicity. *Environ. Mutagen.* 7: 919-921 (1985).
156. Ward, J. M., Griesemer, R. A., and Weisburger, E. K. The mouse liver tumor as an endpoint in carcinogenesis tests. *Toxicol. Appl. Pharmacol.* 51: 389-397 (1979).
157. Reddy, J. K., and Azarnoff, D. L. Hypolipidaemic hepatic peroxisome proliferators form a novel class of chemical carcinogens. *Nature* 283: 397-398 (1980).
158. Maronpot, R. R., Haseman, J. K., Boorman, G. A., Eustis, S. E., Rao, G. N., and Huff, J. E. Liver lesions in B6C3F<sub>1</sub> mice: the National Toxicology Program, experience and position. *Arch. Toxicol. (suppl.)* 10: 10-26 (1987).
159. National Primary Drinking Water Regulations. *Fed. Reg.* 44: 68641 (1979).
160. National Primary Drinking Water Regulations. *Fed. Reg.* 52: 20674 (1987).
161. Proposed Rule. *Fed. Reg.* 39: 42748 (1974).
162. National Academy of Sciences. Health Effects of Chemical Pesticides. Report of the Consultative Panel on Health Hazards of Chemical Pesticides, Vol. 1. National Academy of Sciences Press, Washington, DC, 1975, pp. 67-87.
163. Lyon, J. L., Gardner, J. W., and West, D. W. Cancer incidence in Mormons and non-Mormons in Utah during 1967-75. *J. Natl. Cancer Inst.* 65: 1055-1061 (1980).